

The Ergodic Theory of Hyperbolic Groups

Danny Calegari

ABSTRACT. These notes are a self-contained introduction to the use of dynamical and probabilistic methods in the study of hyperbolic groups. Most of this material is standard; however some of the proofs given are new, and some results are proved in greater generality than have appeared in the literature.

CONTENTS

1. Introduction
 2. Hyperbolic groups
 3. Combings
 4. Random walks
- Acknowledgments
References

1. Introduction

These are notes from a minicourse given at a workshop in Melbourne July 11–15 2011. There is little pretension to originality; the main novelty is firstly that we give a new (and much shorter) proof of Coornaert's theorem on Patterson–Sullivan measures for hyperbolic groups (Theorem 2.5.4), and secondly that we explain how to combine the results of Calegari–Fujiwara in [8] with that of Pollicott–Sharp [35] to prove central limit theorems for quite general classes of functions on hyperbolic groups (Corollary 3.7.5 and Theorem 3.7.6), crucially *without* the hypothesis that the Markov graph encoding an automatic structure is ergodic.

A final section on random walks is much more cursory.

2. Hyperbolic groups

2.1. Coarse geometry. The fundamental idea in geometric group theory is to study groups as automorphisms of geometric spaces, and as a special case, to study the group itself (with its canonical self-action) as a geometric space. This is accomplished most directly by means of the *Cayley graph* construction.

2010 *Mathematics Subject Classification.* Primary 20F10, 20F32, 20F67, 37D20, 60B15, 60J50, 68Q70.

DEFINITION 2.1.1 (Cayley graph). Let G be a group and S a (usually finite) generating set. Associated to G and S we can form the *Cayley graph* $C_S(G)$. This is a graph with vertex set G , and with an edge from g to gs for all $g \in G$ and $s \in S$.

The action of G on itself by (left) multiplication induces a properly discontinuous action of G on $C_S(G)$ by simplicial automorphisms.

If G has no 2-torsion, the action is free and properly discontinuous, and the quotient is a wedge of $|S|$ circles X_S . In this case, if G has a presentation $G = \langle S \mid R \rangle$ we can think of $C_S(G)$ as the covering space of X_S corresponding to the subgroup of the free group F_S normally generated by R , and the action of G on $C_S(G)$ is the deck group of the covering.

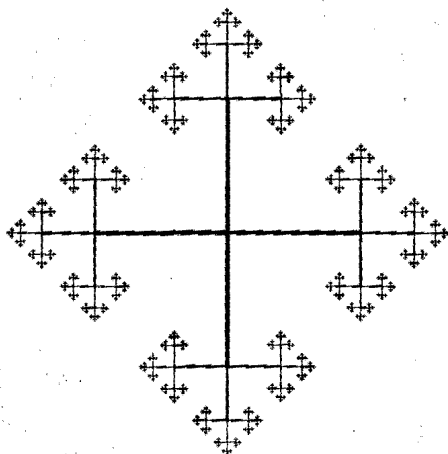


FIGURE 1. The Cayley graph of $F_2 = \langle a, b \mid \rangle$ with generating set $S = \{a, b\}$

We assume the reader is familiar with the notion of a metric space, i.e. a space X together with a symmetric non-negative real-valued function d_X on $X \times X$ which vanishes precisely on the diagonal, and which satisfies the triangle inequality $d_X(x, y) + d_X(y, z) \geq d_X(x, z)$ for each triple $x, y, z \in X$. A metric space is a *path metric space* if for each $x, y \in X$, the distance $d_X(x, y)$ is equal to the infimum of the set of numbers L for which there is a 1-Lipschitz map $\gamma : [0, L] \rightarrow X$ sending 0 to x and L to y . It is a *geodesic metric space* if it is a path metric space and if the infimum is achieved on some γ for each pair x, y ; such a γ is called a *geodesic*. Finally, a metric space is *proper* if closed metric balls of bounded radius are *compact* (equivalently, for each point x the function $d(x, \cdot) : X \rightarrow \mathbb{R}$ is proper).

The graph $C_S(G)$ can be canonically equipped with the structure of a *geodesic metric space*. This is accomplished by making each edge isometric to the Euclidean unit interval. If S is finite, $C_S(G)$ is *proper*. Note that G itself inherits a subspace metric from $C_S(G)$, called the *word metric*. We denote the word metric by d_S , and define $|g|_S$ (or just $|g|$ if S is understood) to be $d_S(\text{id}, g)$. Observe that $d_S(g, h) = |g^{-1}h|_S = |h^{-1}g|_S$ and that $|g|_S$ is the length of the shortest word in elements of S and their inverses representing the element g .

The most serious shortcoming of this construction is its dependence on the choice of a generating set S . Different choices of generating set S give rise to

different spaces $C_S(G)$ which are typically not even homeomorphic. The standard way to resolve this issue is to coarsen the geometric category in which one works.

DEFINITION 2.1.2. Let X, d_X and Y, d_Y be metric spaces. A map $f : X \rightarrow Y$ (not assumed to be continuous) is a *quasi-isometric map* if there are constants $K \geq 1, \epsilon \geq 0$ so that

$$K^{-1}d_X(x_1, x_2) - \epsilon \leq d_Y(f(x_1), f(x_2)) \leq Kd_X(x_1, x_2) + \epsilon$$

for all $x_1, x_2 \in X$. It is said to be a *quasi-isometry* if further $f(X)$ is a *net* in Y ; that is, if there is some R so that Y is equal to the R -neighborhood of $f(X)$.

One also uses the terminology K, ϵ *quasi-isometric map* or K, ϵ *quasi-isometry* if the constants are specified. Note that a $K, 0$ quasi-isometric map is the same thing as a K bilipschitz map. The best constant K is called the *multiplicative constant*, and the best ϵ the *additive constant* of the map.

We denote the R -neighborhood of a set Σ by $N_R(\Sigma)$. Hence a quasi-isometry is a quasi-isometric map for which $Y = N_R(f(X))$ for some R .

REMARK 2.1.3. It is much more common to use the terminology *quasi-isometric embedding* instead of quasi-isometric map as above; we consider this terminology misleading, and therefore avoid it.

LEMMA 2.1.4. *Quasi-isometry is an equivalence relation.*

PROOF. Reflexivity and transitivity are obvious, so we must show symmetry. For each $y \in Y$ choose $x \in X$ with $d_Y(y, f(x)) \leq R$ (such an x exists by definition) and define $g(y) = x$. Observe $d_Y(y, fg(y)) \leq R$ by definition. Then

$$d_X(g(y_1), g(y_2)) \leq Kd_Y(fg(y_1), fg(y_2)) + K\epsilon \leq Kd_Y(y_1, y_2) + K(\epsilon + 2R)$$

Similarly,

$$d_X(g(y_1), g(y_2)) \geq K^{-1}d_Y(fg(y_1), fg(y_2)) - K^{-1}\epsilon \geq K^{-1}d_Y(y_1, y_2) - K^{-1}(\epsilon + 2R)$$

proving symmetry. \square

Note that the compositions fg and gf as above move points a bounded distance. One can define a category in which objects are equivalence classes of metric spaces under the equivalence relation generated by thickening (i.e. isometric inclusion as a net in a bigger space), and morphisms are equivalence classes of quasi-isometric maps, where two maps are equivalent if their values on each point are a uniformly bounded distance apart. In this category, quasi-isometries are isomorphisms. In particular, the set of quasi-isometries of a metric space X , modulo maps that move points a bounded distance, is a *group*, denoted $QI(X)$, which only depends on the quasi-isometry type of X . Determining $QI(X)$, even for very simple spaces, is typically extraordinarily difficult.

EXAMPLE 2.1.5. A metric space X, d_X is quasi-isometric to a point if and only if it has bounded diameter. A Cayley graph $C_S(G)$ (for S finite) is quasi-isometric to a point if and only if G is finite.

EXAMPLE 2.1.6. If S and T are two finite generating sets for a group G then the identity map from G to itself is a quasi-isometry (in fact, a bilipschitz map) of G, d_S to G, d_T . For, there are constants C_1 and C_2 so that $d_T(s) \leq C_1$ for all $s \in S$, and $d_S(t) \leq C_2$ for all $t \in T$, and therefore $C_2^{-1}d_T(g, h) \leq d_S(g, h) \leq C_1d_T(g, h)$.

Because of this, the quasi-isometry class of G , d_S is *independent* of the choice of finite generating set, and we can speak unambiguously of the quasi-isometry class of G .

The Schwarz Lemma connects the geometry of groups to the geometry of spaces they act on.

LEMMA 2.1.7 (Schwarz Lemma). *Let G act properly discontinuously and cocompactly by isometries on a proper geodesic metric space X . Then G is finitely generated by some set S , and the orbit map $G \rightarrow X$ sending g to gx (for any $x \in X$) is a quasi-isometry from G, d_S to X .*

PROOF. Since X is proper and G acts cocompactly there is an R so that $GN_R(x) = X$. Note that Gx is a net, since every point of X is contained in some translate gB and is therefore within distance R of gx .

Let $B = N_{2R+1}(x)$. Since G acts properly discontinuously, there are only finitely many g in G for which $gB \cap B$ is nonempty; let S be the nontrivial elements of this set.

Now, if $g, h \in G$ are arbitrary, let γ be a geodesic in X from gx to hx . Parameterize γ by arclength, and for each integer $i \in (0, |\gamma|)$ let g_i be such that $d_X(g_i x, \gamma(i)) \leq R$. Then $g_i^{-1} g_{i+1} \in S$ and therefore

$$d_S(g, h) = |g^{-1}h| \leq |\gamma| + 1 = d(gx, hx) + 1$$

which shows incidentally that S generates G .

Conversely, if $L := d_S(g, h)$ and g_i is a sequence of elements with $g_0 = g$ and $g_L = h$ and each $g_i^{-1} g_{i+1} \in S$, then there is a path γ_i from $g_i x$ to $g_{i+1} x$ of length at most $4R + 2$, and the concatenation of these paths certifies that

$$d(gx, hx) \leq (4R + 2)|g^{-1}h| = (4R + 2)d_S(g, h)$$

This completes the proof of the lemma. □

EXAMPLE 2.1.8. If G is a group and H is a subgroup of finite index, then G and H are quasi-isometric (for, both act properly discontinuously and cocompactly on $C_S(G)$). Two groups are said to be *commensurable* if they have isomorphic subgroups of finite index; the same argument shows that commensurable groups are quasi-isometric.

EXAMPLE 2.1.9. Any two regular trees of (finite) valence ≥ 3 are quasi-isometric; for, any such tree admits a cocompact action by a free group of finite rank, and any two free groups of finite rank are commensurable.

EXAMPLE 2.1.10. The set of ends of a geodesic metric space is a quasi-isometry invariant. A famous theorem of Stallings [39] says that a finitely generated group with more than one end splits over a finite subgroup; it follows that the property of splitting over a finite subgroup is a quasi-isometry invariant.

Finiteness of the edge groups (in a splitting) is detected quasi-isometrically by the existence of *separating* compact subsets. Quasi-isometry can further detect the finiteness of the vertex groups, and in particular one observes that a group is quasi-isometric to a free group if and only if it is virtually free.

EXAMPLE 2.1.11. Any two groups that act cocompactly and properly discontinuously on the same space X are quasi-isometric. For example, if M_1, M_2 are closed Riemannian manifolds with isometric universal covers, then $\pi_1(M_1)$ and $\pi_1(M_2)$ are

quasi-isometric. It is easy to produce examples for which the groups in question are not commensurable; for instance, a pair of closed hyperbolic 3-manifolds M_1, M_2 with different invariant trace fields (see [27]).

REMARK 2.1.12. In the geometric group theory literature, Lemma 2.1.7 is often called the “Milnor–Švarc (or Švarc–Milnor) Lemma”; “Švarc” here is in fact the well-known mathematical physicist Albert Schwarz; it is our view that the orthography “Švarc” tends to obscure this. Actually, the content of this Lemma was first observed by Schwarz in the early 50’s and only rediscovered 15 years later by Milnor at a time when the work of Soviet mathematicians was not widely disseminated in the west.

2.2. Hyperbolic spaces. In a geodesic metric space a *geodesic triangle* is just a union of three geodesics joining three points in pairs. If the three points are x, y, z we typically denote the (oriented) geodesics by xy, yz and zx respectively; this notation obscures the possibility that the geodesics in question are not uniquely determined by their endpoints.

DEFINITION 2.2.1. A geodesic metric space X, d_X is δ -hyperbolic if for any geodesic triangle, each side of the triangle is contained in the δ -neighborhood of the union of the other two sides. A metric space is *hyperbolic* if it is δ -hyperbolic for some δ .

One sometimes says that *geodesic triangles are δ -thin*.

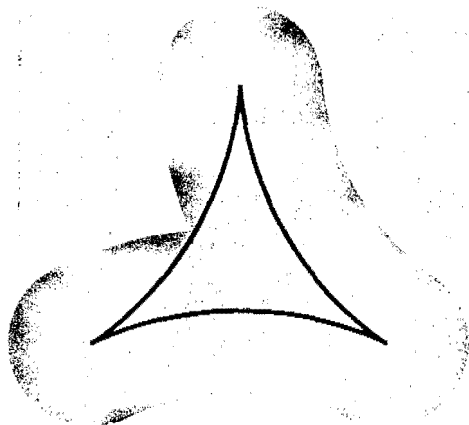


FIGURE 2. A δ -thin triangle; the gray tubes have thickness δ .

EXAMPLE 2.2.2. A tree is 0-hyperbolic.

EXAMPLE 2.2.3. Hyperbolic space (of any dimension) is δ -hyperbolic for a uniform δ .

EXAMPLE 2.2.4. If X is a simply-connected complete Riemannian manifold with curvature bounded above by some $K < 0$ then X is δ -hyperbolic for some δ depending on K .

quasi-isometric. It is easy to produce examples for which the groups in question are not commensurable; for instance, a pair of closed hyperbolic 3-manifolds M_1 , M_2 with different invariant trace fields (see [27]).

REMARK 2.1.12. In the geometric group theory literature, Lemma 2.1.7 is often called the “Milnor–Švarc (or Švarc–Milnor) Lemma”; “Švarc” here is in fact the well-known mathematical physicist Albert Schwarz; it is our view that the orthography “Švarc” tends to obscure this. Actually, the content of this Lemma was first observed by Schwarz in the early 50’s and only rediscovered 15 years later by Milnor at a time when the work of Soviet mathematicians was not widely disseminated in the west.

2.2. Hyperbolic spaces. In a geodesic metric space a *geodesic triangle* is just a union of three geodesics joining three points in pairs. If the three points are x, y, z we typically denote the (oriented) geodesics by xy , yz and zx respectively; this notation obscures the possibility that the geodesics in question are not uniquely determined by their endpoints.

DEFINITION 2.2.1. A geodesic metric space X, d_X is δ -hyperbolic if for any geodesic triangle, each side of the triangle is contained in the δ -neighborhood of the union of the other two sides. A metric space is *hyperbolic* if it is δ -hyperbolic for some δ .

One sometimes says that *geodesic triangles are δ -thin*.

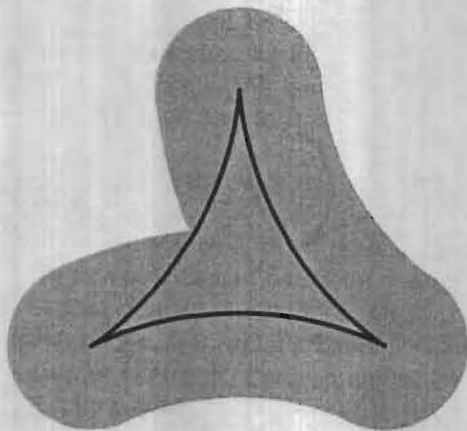


FIGURE 2. A δ -thin triangle; the gray tubes have thickness δ .

EXAMPLE 2.2.2. A tree is 0-hyperbolic.

EXAMPLE 2.2.3. Hyperbolic space (of any dimension) is δ -hyperbolic for a uniform δ .

EXAMPLE 2.2.4. If X is a simply-connected complete Riemannian manifold with curvature bounded above by some $K < 0$ then X is δ -hyperbolic for some δ depending on K .

DEFINITION 2.2.5. A geodesic metric space X is $\text{CAT}(K)$ for some K if triangles are *thinner* than comparison triangles in a space of constant curvature K . This means that if xyz is a geodesic triangle in X , and $x'y'z'$ is a geodesic triangle in a complete simply connected Riemannian manifold Y of constant curvature K with edges of the same lengths, and $\phi : xyz \rightarrow x'y'z'$ is an isometry on each edge, then for any $w \in yz$ we have $d_X(x, w) \leq d_Y(x', \phi(w))$.

The initials CAT stand for Cartan–Alexandrov–Toponogov, who made substantial contributions to the theory of comparison geometry.

EXAMPLE 2.2.6. From the definition, a $\text{CAT}(K)$ space is δ -hyperbolic whenever the complete simply connected Riemannian 2-manifold of constant curvature K is δ -hyperbolic. Hence a $\text{CAT}(K)$ space is hyperbolic if $K < 0$.

EXAMPLE 2.2.7. Nearest point projection to a convex subset of a $\text{CAT}(K)$ space with $K \leq 0$ is distance nonincreasing. Therefore the subspace metric and the path metric on a convex subset of a $\text{CAT}(K)$ space agree, and such a subspace is itself $\text{CAT}(K)$.

Thinness of triangles implies thinness of arbitrary polygons.

EXAMPLE 2.2.8. Let X be δ -hyperbolic and let $abcd$ be a geodesic quadrilateral. Then either there are points on ab and cd at distance $\leq 2\delta$ or there are points on ad and bc at distance $\leq 2\delta$, or possibly both.

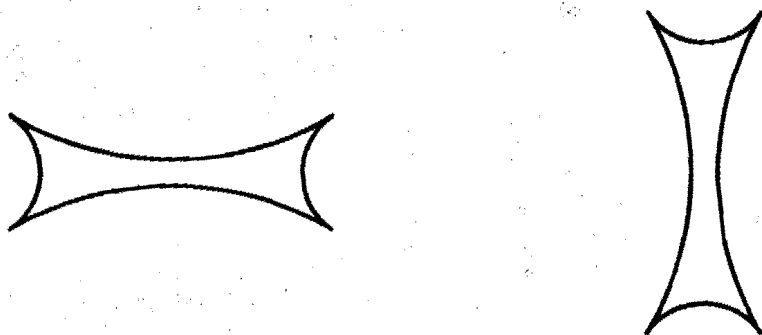


FIGURE 3. Two ways that a quadrilateral can be thin

The number of essentially distinct ways in which an n -gon can be thin is equal to the n th Catalan number. By cutting up a polygon into triangles and examining the implications of δ -thinness for each triangle, one can reason about the geometry of complicated configurations in δ -hyperbolic space.

LEMMA 2.2.9. Let X be δ -hyperbolic, let γ be a geodesic segment/ray/line in X , and let $p \in X$. Then there is a point q on γ realizing the infimum of distance from p to points on γ , and moreover for any two such points q, q' we have $d_X(q, q') \leq 4\delta$.

PROOF. The existence of some point realizing the infimum follows from the properness of $d(p, \cdot) : \gamma \rightarrow \mathbb{R}$, valid for any geodesic in any metric space.

Let q, q' be two such points, and if $d(q, q') > 4\delta$ let q'' be the midpoint of the segment qq' , so $d(q, q'') = d(q'', q') > 2\delta$. Without loss of generality there is r on pq with $d(r, q'') \leq \delta$ hence $d(r, q) > \delta$. But then

$$d(p, q'') \leq d(p, r) + d(r, q'') \leq d(p, r) + \delta < d(p, r) + d(r, q) = d(p, q)$$

contrary to the fact that q minimizes the distance from p to points on γ . \square

Lemma 2.2.9 says that there is an approximate "nearest point projection" map π from X to any geodesic γ (compare with Example 2.2.7). This map is not continuous, but nearby points must map to nearby points, in the sense that $d(\pi(x), \pi(y)) \leq d(x, y) + 8\delta$.

We would now like to show that the property of being hyperbolic is preserved under quasi-isometry. The problem is that the property of δ -hyperbolicity is expressed in terms of geodesics, and quasi-isometries do not take geodesics to geodesics.

A *quasigeodesic segment/ray/line* is the image of a segment/ray/line in \mathbb{R} under a quasi-isometric map. For infinite or semi-infinite intervals this definition has content; for finite intervals this definition has no content without specifying the constants involved. Hence we can talk about a K, ϵ quasigeodesic segment/ray/line.

LEMMA 2.2.10 (Morse lemma). *Let X, d_X be a proper δ -hyperbolic space. Then for any K, ϵ there is a constant C (depending in an explicit way on K, ϵ, δ) so that any K, ϵ quasigeodesic γ is within Hausdorff distance C of a genuine geodesic γ^g . If γ has one or two endpoints, γ^g can be chosen to have the same endpoints.*

PROOF. If γ is noncompact, it can be approximated on compact subsets by finite segments γ_i . If we prove the lemma for finite segments, then a subsequence of the γ_i^g , converging on compact subsets, will limit to γ^g with the desired properties (here is where we use properness of X). So it suffices to prove the lemma for γ a segment.

In this case choose any γ^g with the same endpoints as γ . We need to estimate the Hausdorff distance from γ to γ^g . Fix some constant C and suppose there are points p, p' on γ that are both distance C from γ^g , but $d(r, \gamma^g) \geq C$ for all r on γ between p and p' . Choose p_i a sequence of points on γ and q_i a sequence of points on γ^g closest to the p_i so that $d(q_i, q_{i+1}) = 11\delta$.

Consider the quadrilateral $p_i p_{i+1} q_{i+1} q_i$. By Example 2.2.8 either there are close points on $p_i p_{i+1}$ and $q_i q_{i+1}$, or close points on $p_i q_i$ and $p_{i+1} q_{i+1}$ (or possibly both). Suppose there are points r_i on $p_i q_i$ and r_{i+1} on $p_{i+1} q_{i+1}$ with $d(r_i, r_{i+1}) \leq 2\delta$. Then any nearest point projections of r_i and r_{i+1} to γ^g must be at most distance 10δ apart. But q_i and q_{i+1} are such nearest point projections, by definition, and satisfy $d(q_i, q_{i+1}) = 11\delta$. So it must be instead that there are points r_i on $p_i p_{i+1}$ and s_i on $q_i q_{i+1}$ which are at most 2δ apart. But this means that $d(p_i, p_{i+1}) \geq 2C - 4\delta$, so the length of γ between p and p' is at least $(2C - 4)d(q, q')/11\delta$ where q, q' are points on γ closest to p, p' . On the other hand, $d(p, p') \leq 2C + d(q, q')$. Since γ is a K, ϵ quasigeodesic, if $d(q, q')$ is big enough, we get a uniform bound on C in terms of K, ϵ, δ . The remaining case where $d(q, q')$ is itself uniformly bounded but C is unbounded quickly leads to a contradiction. \square

COROLLARY 2.2.11. *Let Y be δ -hyperbolic and let $f : X \rightarrow Y$ be a K, ϵ quasi-isometry. Then X is δ' -hyperbolic for some δ . Hence the property of being hyperbolic is a quasi-isometry invariant.*

PROOF. Let Γ be a geodesic triangle in X with vertices a, b, c . Then the edges of $f(\Gamma)$ are K, ϵ quasigeodesics in Y , and are therefore within Hausdorff distance C of geodesics with the same endpoints. It follows that every point on $f(ab)$ is within distance $2C + \delta$ of $f(ac) \cup f(bc)$ and therefore every point on ab is within distance $K(2C + \delta) + \epsilon$ of $ac \cup bc$. \square

The Morse Lemma lets us promote quasigeodesics to (nearby) geodesics. The next lemma says that quasigeodesicity is a *local* condition.

DEFINITION 2.2.12. A path γ in X is a *k-local geodesic* if the subsegments of length $\leq k$ are geodesics. Similarly, γ is a *k-local K, ϵ quasigeodesic* if the subsegments of length $\leq k$ are K, ϵ quasigeodesics.

LEMMA 2.2.13 (*k-local geodesics*). *Let X be a δ -hyperbolic geodesic space, and let $k > 8\delta$. Then any k -local geodesic is K, ϵ quasigeodesic for K, ϵ depending explicitly on δ .*

More generally, for any K, ϵ there is a k and constants K', ϵ' so that any k -local K, ϵ quasigeodesic is a K', ϵ' quasigeodesic.

PROOF. Let γ be a k -local geodesic segment from p to q , and let γ^g be any geodesic from p to q . Let r be a point on γ furthest from γ^g , and let r be the midpoint of an arc $r'r''$ of γ of length 8δ . By hypothesis, $r'r''$ is actually a geodesic. Let s' and s'' be points on γ^g closest to r' and r'' . The point r is within distance 2δ either of γ^g or of one of the sides $r's'$ or $r''s''$. If the latter, we would get a path from r to s' or s'' shorter than the distance from r' or r'' , contrary to the definition of r . Hence the distance from r to γ^g is at most 2δ , and therefore γ is contained in the 2δ neighborhood of γ^g .

Now let $\pi : \gamma \rightarrow \gamma^g$ take points on γ to closest points on γ^g . Since π moves points at most 2δ , it is approximately continuous. Since γ is a k -local geodesic, the map π is approximately monotone; i.e. if p_i are points on γ with $d(p_i, p_{i+1}) = k$ moving monotonely from one end of γ to the other, then $d(\pi(p_i), \pi(p_{i+1})) \geq k - 4\delta$ and the projections also move monotonely along γ . In particular, $d(p_i, p_j) \geq (k - 4\delta)|i - j|$ and π is a quasi-isometry. The constants involved evidently depend only on δ and k , and the multiplicative constant evidently goes to 1 as k gets large.

The more general fact is proved similarly, by using Lemma 2.2.10 to promote local quasigeodesics to local geodesics, and then back to global quasigeodesics. \square

2.3. Hyperbolic groups. Corollary 2.2.11 justifies the following definition:

DEFINITION 2.3.1. A group G is hyperbolic if $C_S(G)$ is δ -hyperbolic for some δ for some (and hence for any) finite generating set S .

EXAMPLE 2.3.2. Free groups are hyperbolic, since their Cayley graphs (with respect to a free generating set) are trees which are 0-hyperbolic.

EXAMPLE 2.3.3. Virtually free groups, being precisely the groups quasi-isometric to trees, are hyperbolic. A group quasi-isometric to a point or to \mathbb{R} is finite or virtually \mathbb{Z} respectively; such groups are called *elementary hyperbolic groups*; all others are *nonelementary*.

EXAMPLE 2.3.4. Fundamental groups of closed surfaces with negative Euler characteristic are hyperbolic. By the uniformization theorem, each such surface can be given a hyperbolic metric, exhibiting π_1 as a cocompact group of isometries of the hyperbolic plane.

EXAMPLE 2.3.5. A *Kleinian group* is a finitely generated discrete subgroup of the group of isometries of hyperbolic 3-space. A Kleinian group G is *convex cocompact* if it acts cocompactly on the convex hull of its limit set (in the sphere at infinity). Such a convex hull is CAT(-1), so a convex cocompact Kleinian group is hyperbolic. See e.g. [28] for an introduction to Kleinian groups.

LEMMA 2.3.6 (invariant quasixis). *Let G be hyperbolic. Then there are finitely many conjugacy classes of torsion elements (and therefore a bound on the order of the torsion) and there are constants K, ϵ so that for any nontorsion element g there is a K, ϵ quasigeodesic γ invariant under g on which g acts as translation.*

PROOF. Let $g \in G$ be given. Consider the action of g on the Cayley graph $C_S(G)$. The action is simplicial, so $p \rightarrow d(p, gp)$ has no strict local minima in the interior of edges, and takes integer values at the vertices (which correspond to elements of G). It follows that there is some h for which $d(h, gh)$ is minimal, and we can take h to be an element of G (i.e. a vertex). If $d(h, gh) = k > 8\delta$ then we can join h to gh by a geodesic σ and let $\gamma = \cup_i g^i \sigma$. Note that g acts on γ by translation through distance k ; since this is the minimum distance that g moves points of G , it follows that γ is a k -local geodesic (and therefore a K, ϵ quasigeodesic by Lemma 2.2.13). Note in this case that g has infinite order.

Otherwise there is h moved a least distance by g so that $d(h, gh) \leq 8\delta$. Since G acts cocompactly on itself, there are only finitely many conjugacy classes of elements that move some point any uniformly bounded distance, so if g is torsion we are done. If g is not torsion, its orbits are proper, so for any T there is an N so that $d(h, g^N h) > T$; choose T (and N) much bigger than some fixed (but big) n . Let γ be a geodesic from h to $g^N h$. Then for any $0 \leq i \leq n$ the geodesic $g^i \gamma$ has endpoints within distance $8\delta n$ of the endpoints of γ . On the other hand, $|\gamma| = T \gg 8\delta n$ so γ contains a segment σ of length at least $T - 16\delta n - O(\delta)$ such that $g^i \sigma$ is contained in the 2δ neighborhood of γ for $0 \leq i \leq n$. To see this, consider the quadrilateral with successive vertices $h, g^N h, g^{i+N} h$ and $g^i h$. Two nonadjacent sides must contain points which are at most 2δ apart. Since $N \gg i$, the sides must be γ and $g^i \gamma$. We find σ and $g^i \sigma$ in the region where these two geodesics are close.

Consequently, for any $p \in \sigma$ the sequence $p, gp, \dots, g^n p$ is a K, ϵ quasigeodesic for some uniform K, ϵ independent of n . In particular there is a constant C (independent of n) so that $d(p, g^i p) \geq iC$ for $0 \leq i \leq n$, and therefore the infinite sequence $g^i p$ for $i \in \mathbb{Z}$ is an (nC) -local K, ϵ quasigeodesic. Since K, ϵ is fixed, if n is big enough, this infinite sequence is an honest K', ϵ' quasigeodesic invariant under g , by Lemma 2.2.13. Here K', ϵ' depends only on δ and G , and not on g . \square

Lemma 2.3.6 can be weakened considerably, and it is frequently important to study actions which are not necessarily cocompact on δ -hyperbolic spaces which are not necessarily proper. The quasigeodesic γ invariant under g is called a *quasixis*. Quasixes in δ -hyperbolic spaces are (approximately) *unique*:

LEMMA 2.3.7. *Let G be hyperbolic, and let g have infinite order. Let γ and γ' be g -invariant K, ϵ quasigeodesics (i.e. quasixes for g). Then γ and γ' are a finite Hausdorff distance apart, and this finite distance depends only on K, ϵ and δ . Consequently the centralizer $C(g)$ is virtually \mathbb{Z} .*

PROOF. Let $p \in \gamma$ and $p' \in \gamma'$ a closest point to p . Since g acts on both γ and γ' cocompactly, there is a constant C so that every point in γ or γ' is within C

from some point in the orbit of p or p' . This implies that the Hausdorff distance from γ to γ' is at most $2C + d(p, p')$; in particular, this distance is finite.

Pick two points on γ very far away from each other; each is distance at most $2C + d(p, p')$ from γ' , and therefore most of the geodesic between them is within distance 2δ of the geodesic between corresponding points on γ' . But γ and γ' are themselves K, ϵ quasigeodesic, and therefore uniformly close to these geodesics. Hence some points on γ are within a uniformly bounded distance of γ' , and therefore all points on γ are.

If h commutes with g , then h must permute the quiaxes of g . Therefore h takes points on any quiaxis γ for g to within a bounded distance of γ . Hence $C(g)$, thought of as a subset of G , is quasiisometric to a quiaxis (that is to say, to \mathbb{R}), and is therefore virtually \mathbb{Z} . \square

This shows that a hyperbolic group cannot contain a copy of $\mathbb{Z} \oplus \mathbb{Z}$ (or, for that matter, the fundamental group of a Klein bottle). This is more subtle than it might seem; $\mathbb{Z} \oplus \mathbb{Z}$ can act freely and properly discontinuously by isometries on a proper δ -hyperbolic space — for example, as a parabolic subgroup of the isometries of \mathbb{H}^3 .

EXAMPLE 2.3.8. If M is a closed 3-manifold, then $\pi_1(M)$ is hyperbolic if and only if it does not contain any $\mathbb{Z} \oplus \mathbb{Z}$ subgroup. Note that this includes the possibility that $\pi_1(M)$ is elementary hyperbolic (for instance, finite). This follows from Perelman's Geometrization Theorem [31, 32].

If g is an isometry of any metric space X , the *translation length* of g is the limit $\tau(g) := \lim_{n \rightarrow \infty} d_X(p, g^n p)/n$ for some $p \in X$. The triangle inequality implies that the limit exists and is independent of the choice of p . Moreover, from the definition, $\tau(g^n) = |n|\tau(g)$ and $\tau(g)$ is a conjugacy invariant.

Lemma 2.3.6 implies that for G acting on itself, $\tau(g) = 0$ if and only if g has finite (and therefore bounded) order. Consequently a hyperbolic group cannot contain a copy of a Baumslag-Solitar group; i.e. a group of the form $BS(p, q) := \langle a, b \mid ba^p b^{-1} = a^q \rangle$. For, we have already shown hyperbolic groups do not contain $\mathbb{Z} \oplus \mathbb{Z}$, and this rules out the case $|p| = |q|$, and if $|p| \neq |q|$ then for any isometric action of $BS(p, q)$ on a metric space, $\tau(a) = 0$.

By properness of $C_S(G)$ and the Morse Lemma, there is a constant N so that for any $g \in G$ the power g^N has an invariant *geodesic* axis on which it acts by translation. It follows that $\tau(g) \in \mathbb{Q}$, and in fact $\in \frac{1}{N}\mathbb{Z}$; this cute observation is due to Gromov [20].

2.4. The Gromov boundary. Two geodesic rays γ, γ' in a metric space X are *asymptotic* if they are a finite Hausdorff distance apart. The property of being asymptotic is an equivalence relation, and the set of equivalence classes is the *Gromov boundary*, and denoted $\partial_\infty X$. If X is proper and δ -hyperbolic, and x is any basepoint, then every equivalence class contains a ray starting at x . For, if γ is a geodesic ray, and $g_i \in \gamma$ goes to infinity, then by properness, any collection of geodesics xg_i contains a subsequence which converges on compact subsets to a ray γ' . By δ -thinness each of the triangles xg_0g_i is contained in a uniformly bounded neighborhood of γ , so the same is true of γ' ; in particular, γ' is asymptotic to γ . We give $\partial_\infty X$ the topology of convergence on compact subsets of equivalence classes. That is, $\gamma_i \rightarrow \gamma$ if and only if every subsequence of the γ_i contains a

further subsequence whose equivalence classes have representatives that converge on compact subsets to some representative of the equivalence class of γ .

LEMMA 2.4.1. *Let X be a δ -hyperbolic proper geodesic metric space. Then $\partial_\infty X$ is compact.*

PROOF. If γ_i is any sequence of rays, and γ'_i is an equivalent sequence starting at a basepoint x , then by properness γ'_i has a subsequence which converges on compact subsets. \square

In fact, we can define a (compact) topology on $\overline{X} := X \cup \partial_\infty X$ by saying that $x_i \rightarrow \gamma$ if and only if every subsequence of a sequence of geodesics xx_i contains a further subsequence which converges on compact subsets to a representative of γ . With this topology, \overline{X} is compact, $\partial_\infty X$ is closed in \overline{X} , and the inclusion of X into \overline{X} is a homeomorphism onto its image.

A bi-infinite geodesic γ determines two (distinct) points in $\partial_\infty X$; we call these the *endpoints* of γ . Two geodesics with the same (finite or infinite) endpoints are Hausdorff distance at most 2δ apart. Conversely, any two distinct points in $\partial_\infty X$ are spanned by an infinite geodesic γ . For, if γ_1, γ_2 are two infinite rays (starting at a basepoint x for concreteness), and g_i, h_i are points on γ_1, γ_2 respectively going to infinity, some point p_i on any geodesic $g_i h_i$ is within δ of both xg_i and xh_i , and if $p_i \rightarrow \infty$ then γ_1 and γ_2 would be a finite Hausdorff distance apart. Otherwise some subsequence of the p_i converges to p , and the geodesics $g_i h_i$ converge on compact subsets to a (nonempty!) bi-infinite geodesic γ through p asymptotic to both γ_1 and γ_2 . Evidently, geodesic triangles with some or all endpoints at infinity are δ' -thin for some δ' depending only on δ (one can take $\delta' = 20\delta$). By abuse of notation, in the sequel we will call a metric space δ -hyperbolic if all geodesic triangles — even those with some endpoints at infinity — are δ -thin.

Let X, Y be hyperbolic geodesic metric spaces. Then any quasi-isometric map $\phi : X \rightarrow Y$ extends uniquely to a *continuous* map $\partial_\infty X \rightarrow \partial_\infty Y$. In particular, the Gromov boundary $\partial_\infty X$ depends (up to homeomorphism) only on the quasi-isometry type of X , and $QI(X)$ acts on $\partial_\infty X$ by homeomorphisms.

If G is a hyperbolic group, we define $\partial_\infty G$ to be the Gromov boundary of some (any) $C_S(G)$.

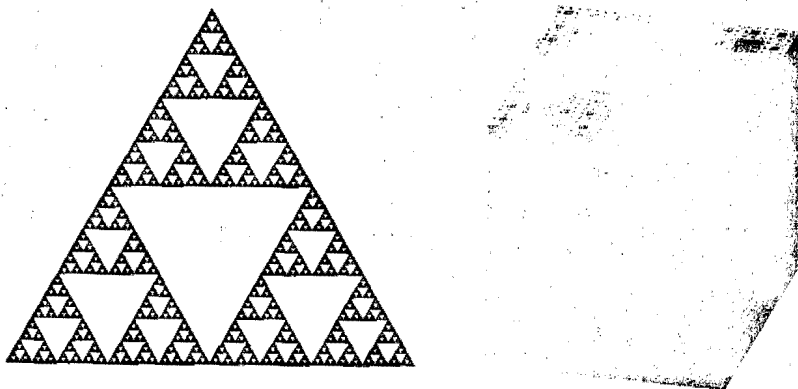


FIGURE 4. The Sierpinski carpet and the Menger sponge.

further subsequence whose equivalence classes have representatives that converge on compact subsets to some representative of the equivalence class of γ .

LEMMA 2.4.1. *Let X be a δ -hyperbolic proper geodesic metric space. Then $\partial_\infty X$ is compact.*

PROOF. If γ_i is any sequence of rays, and γ'_i is an equivalent sequence starting at a basepoint x , then by properness γ'_i has a subsequence which converges on compact subsets. \square

In fact, we can define a (compact) topology on $\bar{X} := X \cup \partial_\infty X$ by saying that $x_i \rightarrow \gamma$ if and only if every subsequence of a sequence of geodesics xx_i contains a further subsequence which converges on compact subsets to a representative of γ . With this topology, \bar{X} is compact, $\partial_\infty X$ is closed in \bar{X} , and the inclusion of X into \bar{X} is a homeomorphism onto its image.

A bi-infinite geodesic γ determines two (distinct) points in $\partial_\infty X$; we call these the *endpoints* of γ . Two geodesics with the same (finite or infinite) endpoints are Hausdorff distance at most 2δ apart. Conversely, any two distinct points in $\partial_\infty X$ are spanned by an infinite geodesic γ . For, if γ_1, γ_2 are two infinite rays (starting at a basepoint x for concreteness), and g_i, h_i are points on γ_1, γ_2 respectively going to infinity, some point p_i on any geodesic $g_i h_i$ is within δ of both xg_i and xh_i , and if $p_i \rightarrow \infty$ then γ_1 and γ_2 would be a finite Hausdorff distance apart. Otherwise some subsequence of the p_i converges to p , and the geodesics $g_i h_i$ converge on compact subsets to a (nonempty!) bi-infinite geodesic γ through p asymptotic to both γ_1 and γ_2 . Evidently, geodesic triangles with some or all endpoints at infinity are δ' -thin for some δ' depending only on δ (one can take $\delta' = 20\delta$). By abuse of notation, in the sequel we will call a metric space δ -hyperbolic if all geodesic triangles — even those with some endpoints at infinity — are δ -thin.

Let X, Y be hyperbolic geodesic metric spaces. Then any quasi-isometric map $\phi : X \rightarrow Y$ extends uniquely to a continuous map $\partial_\infty X \rightarrow \partial_\infty Y$. In particular, the Gromov boundary $\partial_\infty X$ depends (up to homeomorphism) only on the quasi-isometry type of X , and $\text{QI}(X)$ acts on $\partial_\infty X$ by homeomorphisms.

If G is a hyperbolic group, we define $\partial_\infty G$ to be the Gromov boundary of some (any) $C_S(G)$.

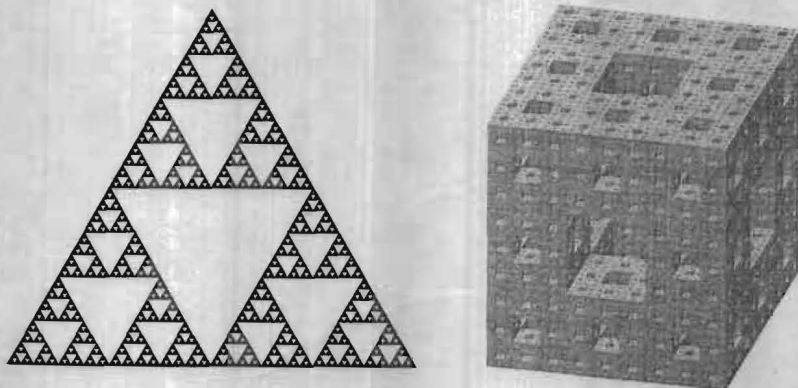


FIGURE 4. The Sierpinski carpet and the Menger sponge.

EXAMPLE 2.4.2. If G is free, $\partial_\infty G$ is a Cantor set. If G is a π_1 of a closed surface with negative Euler characteristic, $\partial_\infty G$ is a circle. If G is a convex cocompact Kleinian group, $\partial_\infty G$ is homeomorphic to the limit set. For example, if G is the fundamental group of a hyperbolic 3-manifold with totally geodesic boundary, $\partial_\infty G$ is a Sierpinski carpet.

In fact, a theorem of Kapovich–Kleiner [25] says that if G is a hyperbolic group which does not split over a finite or virtually cyclic subgroup, and if $\partial_\infty G$ is 1-dimensional (in the topological sense of dimension), then $\partial_\infty G$ is homeomorphic to the circle, the Sierpinski carpet, or the Menger sponge.

Evidently, $\partial_\infty G$ is empty if and only if G is finite, and if $\partial_\infty G$ is nonempty, it has at least two points, and has exactly two points if and only if G is itself quasi-isometric to the geodesic joining these two points, which holds if and only if G is virtually \mathbb{Z} .

If $g \in G$ has infinite order, a *quasi-axis* γ is *asymptotic* to two points $p^\pm \in \partial_\infty G$. Under (positive) powers of g , points stay a constant distance from γ , and move towards one of the endpoints, say p^+ . As homeomorphisms from \bar{X} to itself, the elements g^n with $n \rightarrow \infty$ converge uniformly (in the compact-open topology) on $\bar{X} - p^-$ to the constant map to p^+ . We call p^+ the *attracting endpoint* and p^- the *repelling endpoint* of g ; the actions of g on $\partial_\infty G$ is sometimes expressed by saying that it has *source-sink dynamics*.

EXAMPLE 2.4.3. Let $g, h \in G$ be of infinite order, with quasi-axes γ and γ' . If γ and γ' share an endpoint (without loss of generality the attracting endpoint of each) and p is close to both γ and γ' , then there are $n_i, m_i \rightarrow \infty$ for which $d(h^{-m_i} g^{n_i} p, p)$ is bounded. Since the action of G on its Cayley graph is properly discontinuously, it follows that there are distinct i, j with $h^{-m_i} g^{n_i} = h^{-m_j} g^{n_j}$ so that $h^m = g^n$ for some positive n, m . In particular, in this case g and h together generate a virtual \mathbb{Z} subgroup, and their quasi-axes have the same endpoints. Otherwise the endpoints are disjoint, and because of the source-sink dynamics, Klein's pingpong argument implies that sufficiently large powers g^n, h^m generate a (nonabelian) *free* subgroup of G .

LEMMA 2.4.4. *Suppose G is nonelementary. Then the action of G on $\partial_\infty G$ is minimal; i.e. every orbit is dense. Consequently $\partial_\infty G$ is infinite and perfect.*

PROOF. If G is nonelementary, there are g, h whose quasi-axes have distinct endpoints p^\pm and q^\pm respectively. If $r \in \partial_\infty G$ is arbitrary, then either $g^n r \rightarrow p^+$ or $g^n h r \rightarrow p^+$; it follows that every attracting/repelling point is in the closure of every orbit.

Now let γ be a geodesic from p^- to p^+ and let γ' be a geodesic ray asymptotic to r . Pick s on γ and let g_i be a sequence of elements with $g_i(s) \in \gamma'$ converging to r . At most one component of $\gamma - s$ can come close to the basepoint x . Hence there is some subsequence so that either $g_i p^+ \rightarrow r$ or $g_i p^- \rightarrow r$, and therefore every point is in the closure of the orbit of some attracting/repelling point. This proves the lemma. \square

Another way to see the compactification $\partial_\infty X$ is in terms of (equivalence classes of) *horofunctions*.

DEFINITION 2.4.5 (horofunction). Let γ be a geodesic ray parameterized by length. The *horofunction* (also called the *Busemann function*) associated to γ is

the limit

$$b_\gamma(x) := \lim_{t \rightarrow \infty} d_X(x, \gamma(t)) - t$$

The level sets of horofunctions are called *horospheres*.

This limit exists and is finite, by the triangle inequality. Moreover, it is 1-Lipschitz. If γ and γ' are asymptotic, then there is some constant $C(\gamma, \gamma')$ so that $|(b_\gamma - b_{\gamma'}) - C| \leq 2\delta$. If x is the endpoint of γ , we let b_x denote any horofunction of the form b_γ , and say that b_x is *centered at x* .

Here is another way to define b_γ without reference to γ . On any proper metric space, the set of 1-Lipschitz functions mod constants is compact (in the topology of convergence on compact subsets). For any $x \in X$ the function $d_X(x, \cdot) : X \rightarrow \mathbb{R}$ is 1-Lipschitz, and $x \rightarrow d_X(x, \cdot)$ embeds X in the space of 1-Lipschitz functions on X mod constants. The closure of this image defines a natural compactification of X ; quotienting further by bounded functions gives \bar{X} . For each $x \in \partial_\infty X$ the preimage is the set of equivalence classes of functions b_x . In this way we think of b_x as a normalization of the function which measures "distance to x ".

The space $\partial_\infty X$ can be metrized following Gromov (see [20]).

DEFINITION 2.4.6. Fix some basepoint x and some constant $a > 1$. The *a-length* of a rectifiable path γ in X is the integral along γ of $a^{-d_X(x, \cdot)}$, and the *a-distance* from y to z , denoted $d_X^a(y, z)$, is the infimum of the *a-lengths* of paths between y and z .

A straightforward calculation shows that there is an $a_0 > 1$ so that for $a < a_0$, the *a-length* defines a metric on \bar{X} . In fact, any a_0 with $\delta \log(a_0) \ll 1$ will work. If a is too big, *a-length* still extends to a pseudo-metric on \bar{X} , but now distinct points of $\partial_\infty X$ might be joined by a sequence of paths with *a-length* going to 0. Increasing a will decrease the Hausdorff dimension of $\partial_\infty X$; of course, the Hausdorff dimension must always be at least as big as the topological dimension. In any case, it follows that $\partial_\infty X$ is metrizable.

The following lemma is useful to compare length and *a-length*.

LEMMA 2.4.7. For $a < a_0$ there is a constant λ so that for all points $y, z \in \partial_\infty X$ there is an inequality $\lambda^{-1} a^{-d_X(x, yz)} \leq d_X^a(y, z) \leq \lambda a^{-d_X(x, yz)}$ where $d_X(x, yz)$ is the ordinary distance from the basepoint x to the geodesic yz .

For a proof, see [13].

The quantity $d_X(x, yz)$ is sometimes abbreviated by $(y|z)$ (the basepoint x is suppressed in this notation), and called the *Gromov product*. So we can also write $\lambda^{-1} a^{-(y|z)} \leq d_X^a(y, z) \leq \lambda a^{-(y|z)}$. Because of this inequality, different choices of a give rise to Hölder equivalent metrics on $\partial_\infty X$. If X is a group G , we take id as the basepoint, by convention.

REMARK 2.4.8. With our notation, $(y|z) := d_X(x, yz)$ is ambiguous, since it depends on a choice of geodesic from y to z . Since we only care about $(y|z)$ up to a uniform additive constant, we ignore this issue. One common normalization, adopted by Gromov, is to use the formula $(y|z) := \frac{1}{2}(d_X(x, y) + d_X(x, z) - d_X(y, z))$. These definitions are interchangeable for our purposes, as the ambiguity can always be absorbed into some unspecified constant.

A group G acting by homeomorphisms on a compact metrizable space M is said to be a *convergence action* if the induced action on the space $M^3 - \Delta$ of distinct ordered triples is properly discontinuous.

LEMMA 2.4.9. *The action of G on $\partial_\infty G$ is a convergence action. Moreover, the action on the space of distinct triples is cocompact.*

PROOF. If x, y, z is a distinct triple of points in $\partial_\infty G$, there is a point p within distance δ of all three geodesics xy, yz, zx ; moreover, the set of such points has uniformly bounded diameter in G . This defines an approximate map from distinct triples to points in G . Since the action of G on itself is cocompact, the same is true for the action on the space of distinct triples. Similarly, if the action of G on the space of distinct triples were not properly discontinuous, we could find two bounded regions in G and infinitely many g_i in G taking some point in one bounded region to some point in the other, which is absurd. \square

The converse is a famous theorem of Bowditch:

THEOREM 2.4.10 (Bowditch's convergence theorem [3], Thm. 0.1). *Let G act faithfully, properly discontinuously and cocompactly on the space of distinct triples of some perfect compact metrizable space M . Then G is hyperbolic and M is G -equivariantly homeomorphic to $\partial_\infty G$.*

2.5. Patterson-Sullivan measure. The results in this section are due to Coornaert [12], although because of our more narrow focus we are able to give somewhat different and shorter proofs. However by and large our proofs, like Coornaert's, are obtained by directly generalizing ideas of Sullivan [43] in the context of Kleinian groups.

Let G be a hyperbolic group, and let $G_{\leq n}$ denote the set of elements of (word) length $\leq n$, with respect to some fixed generating set. The *critical exponent* $h(G)$ (also called the *volume entropy* of G) is the quantity

$$h(G) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log |G_{\leq n}|$$

in other words, the exponential growth rate of G . Since every nonelementary hyperbolic group contains many free groups (Example 2.4.3), $h(G) = 0$ if and only if G is elementary.

Define the (Poincaré) *zeta function* by the formula

$$\zeta_G(s) := \sum_{g \in G} e^{-s|g|}$$

Then $\zeta_G(s)$ diverges if $s < h(G)$ and converges if $s > h(G)$.

LEMMA 2.5.1. *The zeta function diverges at $s = h(G)$.*

PROOF. We will show in § 3 (Theorem 3.2.2) that for any hyperbolic group G and any generating set S there is a *regular language* $L \subset S^*$ consisting of geodesics, which evaluates bijectively to G . In particular, $|G_{\leq n}| = |L_{\leq n}|$ for any n . In any regular language L the generating function $\sum |L_{\leq n}| t^n$ is *rational* (Theorem 3.1.3); i.e. it is the power series expansion of $p(t)/q(t)$ for some integral polynomials p, q , and consequently $C^{-1}(e^{hn}n^k) \leq |L_{\leq n}| \leq C(e^{hn}n^k)$ for some real h and non-negative integer k , and constant C . Evidently, for L as above, $h = h(G)$ and the zeta function diverges at h . \square

For $s > h(G)$ construct a probability measure ν_s on \bar{G} (i.e. on $G \cup \partial_\infty G$) supported in G , by putting a Dirac mass of size $e^{-s|g|}/\zeta_G(s)$ at each $g \in G$. As s converges to h from above, this sequence of probability measures contains a subsequence which converges to a limit ν . Since the zeta function diverges at h , the limit ν is supported on ∂G . This measure is called a *Patterson-Sullivan measure*, by analogy with the work of Patterson and Sullivan [30, 43] on Kleinian groups.

For any g , the pushforward of measure $g_*\nu_s$ is defined by $g_*\nu_s(A) = \nu_s(g^{-1}A)$, and similarly for $g_*\nu$. For any g, g' there is an inequality $|g'| - |g| \leq |gg'| \leq |g'| + |g|$. From the definition on ν_s , this implies that $g_*\nu_s$ is absolutely continuous with respect to ν_s , and its Radon-Nikodym derivative satisfies $e^{-s|g|} \leq d(g_*\nu_s)/d\nu_s \leq e^{s|g|}$. Passing to a limit we deduce that $e^{-h|g|} \leq d(g_*\nu)/d\nu \leq e^{h|g|}$.

The most important property of the measure ν is a refinement of this inequality, which can be expressed by saying that it is a so-called *quasiconformal measure of dimension h* . The "conformal" structure on $\partial_\infty X$ is defined using the a -distance for some fixed $a > 1$ (recall Definition 2.4.6).

DEFINITION 2.5.2 (Coornaert). For $g \in G$ define $j_g : \partial_\infty X \rightarrow \mathbb{R}$ by

$$j_g(y) = a^{b_y(\text{id}) - b_y(g)}$$

for some horofunction b_y centered at y . A probability measure ν on $\partial_\infty X$ is a *quasiconformal measure of dimension D* if $g_*\nu$ is absolutely continuous with respect to ν for every $g \in G$, and there is some constant C independent of g so that

$$C^{-1}j_g(y)^D \leq d(g_*\nu)/d\nu \leq Cj_g(y)^D$$

Notice that the ambiguity in the choice of horofunction b_y is absorbed into the definition of j_g (which only depends on b_y mod constant functions) and the constant C . The support of any quasiconformal measure is evidently closed and G -invariant, so by Lemma 2.4.4, it is all of $\partial_\infty G$.

From the definition of the Radon-Nikodym derivative, ν is a quasiconformal measure of dimension D if there is a constant C so that for all y we can find a neighborhood V of y in \bar{X} for which

$$C^{-1}j_g(y)^D \nu(A) \leq \nu(g^{-1}A) \leq Cj_g(y)^D \nu(A)$$

for all $A \subset V$.

REMARK 2.5.3. For some reason, Coornaert chooses to work with *pullbacks* of measure $g^*\nu := g_*^{-1}\nu$ instead of pushforward. Therefore the roles of g and g^{-1} are generally interchanged between our discussion and Coornaert's.

THEOREM 2.5.4 (Coornaert [12], Thm. 5.4). *The measure ν is a quasiconformal measure of dimension D where $D = h/\log a$.*

PROOF. Evidently the support of ν is G -invariant, and is therefore equal to all of $\partial_\infty G$. Let $y \in \partial_\infty X$, let b_y be a horofunction centered at y , and let $g \in G$.

By δ -thinness and the definition of a horofunction, $d(g, z) - d(\text{id}, z)$ is close to $b_y(g) - b_y(\text{id})$ for z sufficiently close to y . In particular, there is a neighborhood V of y in \bar{X} so that

$$|g^{-1}z| - |z| - C \leq b_y(g) - b_y(\text{id}) \leq |g^{-1}z| - |z| + C$$

for some C , and for all z in V .

For each $s > h$ we have $g_*\nu_s(z)/\nu_s(z) = \nu_s(g^{-1}z)/\nu_s(z) = e^{-s(|g^{-1}z| - |z|)}$. Taking the limit as $s \rightarrow h$ and defining D by $a^D = e^h$ proves the theorem. \square

To make use of this observation, we introduce the idea of a *shadow*, following Sullivan.

DEFINITION 2.5.5. For $g \in G$ and R a positive real number, the *shadow* $S(g, R)$ is the set of $y \in \partial_\infty G$ such that every geodesic ray from id to y comes within distance R of g .

Said another way, y is in $S(g, R)$ if g comes within distance R of any geodesic from id to y . Given $R > 2\delta$, for any fixed n the shadows $S(g, R)$ with $|g| = n$ cover $\partial_\infty G$ efficiently:

LEMMA 2.5.6. Fix R . Then there is a constant N so that for any $y \in \partial_\infty G$ and any n there is at least 1 and there are at most N elements g with $|g| = n$ and $y \in S(g, R)$.

PROOF. If $R > 2\delta$, if γ is any geodesic from id to y , and if g is any point on γ , then $y \in S(g, R)$. Conversely, if g and h are any two elements with $|g| = |h|$ and $y \in S(g, R) \cap S(h, R)$ then $d(g, h) \leq 2R$. \square

Sullivan's fundamental observation is that the action of g^{-1} on $S(g, R)$ is uniformly close to being *linear*, in the sense that the derivative $d(g_*\nu)/d\nu$ varies by a bounded multiplicative constant on $S(g, R)$:

LEMMA 2.5.7. Fix R . Then there is a constant C so that for any $y \in S(g, R)$ there is an inequality

$$C^{-1}a^{|g|} \leq j_g(y) \leq Ca^{|g|}$$

PROOF. Recall $j_g(y) = a^{b_\nu(\text{id}) - b_\nu(g)}$ for some horofunction b_ν . But by δ -thinness and the definition of a shadow, there is a constant C' so that

$$|g| - C' \leq b_\nu(\text{id}) - b_\nu(g) \leq |g| + C'$$

for any y in $S(g, R)$. \square

From this one readily obtains a uniform estimate on the measure of a shadow:

LEMMA 2.5.8. Fix R . Then there is a constant C so that for any $g \in G$ there is an inequality

$$C^{-1}a^{-|g|D} \leq \nu(S(g, R)) \leq Ca^{-|g|D}$$

PROOF. Let $m_0 < 1$ be the measure of the biggest atom of ν , and fix $m_0 < m < 1$. By compactness of $\partial_\infty G$ there is some ϵ so that every ball in $\partial_\infty G$ of diameter $\leq \epsilon$ has mass at most m . Now, $g^{-1}S(g, R)$ is the set of $y \in \partial_\infty G$ for which every geodesic ray from g^{-1} to y comes within distance R of id . As $R \rightarrow \infty$, the diameter of $\partial_\infty G - g^{-1}S(g, R)$ goes to zero uniformly in g (this follows from the quasi-equivalence of $d_X^q(y, z)$ and $a^{-(y|z)}$; see Lemma 2.4.7). Consequently there is some R_0 so that for all $R \geq R_0$ the measure $\nu(g^{-1}S(g, R))$ is between $1 - m$ and 1, independent of g .

Now, by Lemma 2.5.7 and the definition of a quasiconformal measure, there is a constant C_1 so that

$$C_1^{-1}a^{|g|D} \leq \nu(g^{-1}S(g, R))/\nu(S(g, R)) \leq C_1a^{|g|D}$$

Taking reciprocals, and using the fact that $1 - m \leq \nu(g^{-1}S(g, R)) \leq 1$ completes the proof (at the cost of adjusting constants). \square

Note that the argument shows that ν has no atoms, since any $y \in \partial_\infty G$ is contained in some shadow of measure $\leq Ca^{-Dn}$ for any n . We deduce the following corollary.

COROLLARY 2.5.9 (Coornaert [12], Thm. 7.2). *Let G be a hyperbolic group. Then there is a constant C so that*

$$C^{-1}e^{hn} \leq |G_{\leq n}| \leq Ce^{hn}$$

for all n .

PROOF. The lower bound is proved in § 3, so we just need to prove the upper bound.

For each g with $|g| = n$ Lemma 2.5.8 says $e^{-hn} = a^{-Dn} \leq C_1 \nu(S(g, R))$. On the other hand, Lemma 2.5.6 says that every point $y \in \partial_\infty G$ is in at most N sets $S(g, R)$ with $|g| = n$, so

$$|G_n|e^{-hn}C_1^{-1} \leq \sum_{|g|=n} \nu(S(g, R)) \leq N\nu(\cup_{|g|=n} S(g, R)) = N$$

□

Corollary 2.5.9 has important consequences that we will explore in § 3.

A second corollary gives very precise metric and dynamical control over $\partial_\infty G$. An action of a group G on a space X is said to be *ergodic* for some measure ν on X if for any two subsets A, B of X with $\nu(A), \nu(B) > 0$ there is some $g \in G$ with $\nu(g(A) \cap B) > 0$.

COROLLARY 2.5.10 (Coornaert [12], Cor. 7.5 and Thm. 7.7). *Let ν be a quasiconformal measure on $\partial_\infty G$ of dimension D . Then ν is quasi-equivalent to D -dimensional Hausdorff measure; i.e. there is a constant C so that $C^{-1}H^D(A) \leq \nu(A) \leq CH^D(A)$ for any A . In particular, the space $\partial_\infty G$ has Hausdorff dimension D , and its D -dimensional Hausdorff measure is finite and positive. Moreover, the action of G on ∂G is ergodic for ν .*

PROOF. Evidently, the second and third claims follow from the first (if A is a G -invariant subset of $\partial_\infty G$ of positive ν -measure, the restriction of ν to A is a quasiconformal measure of dimension D , and is therefore quasi-equivalent to H^D and thence to ν . In particular, A has full measure). So it suffices to show that ν and H^D are quasi-equivalent.

Since $C_1^{-1}a^{-(y|z)} \leq d_X^a(y, z) \leq C_1a^{-(y|z)}$ it follows that every metric ball $B(y, r)$ in $\partial_\infty G$ can be sandwiched between two shadows $S(g_1, R) \subset B(y, r) \subset S(g_2, R)$ with $a^{-|g_1|} \geq r/C_2$ and $a^{-|g_2|} \leq rC_2$. From Lemma 2.5.8 we obtain $C_2^{-1}r^D \leq \nu(B(y, r)) \leq C_2r^D$. From this and the definition of Hausdorff measure, we will obtain the theorem.

If A is any measurable set, cover A by balls U_i of radius $\epsilon_i \leq \epsilon$. Then

$$\nu(A) \leq \nu(\cup_i U_i) \leq \sum_i \nu(U_i) \leq C_2 \sum_i \epsilon_i^D$$

so letting $\epsilon \rightarrow 0$ we get $\nu(A) \leq C_2H^D(A)$.

The following proof of the reverse inequality was suggested to us by Curt McMullen. For any δ let K be compact and U open so that $K \subset A \subset U$ and both $\nu(U - K)$ and $H^D(U - K)$ are less than δ . By compactness, there is an ϵ so that

every ball of radius $\leq \epsilon$ centered at a point in K is contained in U . Now inductively cover K by balls U_1, U_2, \dots of non-increasing radius $\epsilon_i \leq \epsilon$ in such a way that the center of each U_i is not in U_j for and $j < i$. Then the balls with the same centers and half the radii are disjoint, so

$$\sum \epsilon_i^D = 2^D \sum (\epsilon_i/2)^D \leq C_3 \nu(U)$$

and therefore $H^D(K) \leq C_3 \nu(U)$. Taking $\delta \rightarrow 0$ gives $H^D(A) \leq C_3 \nu(A)$ and we are done. \square

REMARK 2.5.11. Coornaert only gives the proof of the inequality $\nu(A) \leq CH^D(A)$ in his paper, referring the reader to Sullivan [43] for the proof of $C^{-1}H^D(A) \leq \nu(A)$. However, there is a gap in Sullivan's proof of the reverse inequality, of which the reader should be warned.

REMARK 2.5.12. The approximate linearity of g^{-1} on $S(g, R)$ has many other applications. For example, see the proof of Theorem 1 in [42].

3. Comings

On a Riemannian manifold, a "geodesic" is just a smooth path that locally minimizes length (really, energy). A sufficiently long geodesic is typically not globally length minimizing, and the entire subject of Morse theory is devoted to the difference. By contrast, one of the most important qualitative features of negative curvature is that (quasi)-geodesicity is a *local* property (i.e. Lemma 2.2.13).

This localness translates into an important combinatorial property, known technically as *finiteness of cone types*. This is the basis of Cannon's theory of hyperbolic groups, and for the more general theory of automatic groups and structures (see [15] for more details).

3.1. Regular languages. Let S be a finite set, and let S^* denote the set of finite words in the alphabet S . An *automaton* is a finite directed graph Γ with a distinguished initial vertex, and edges labeled by elements of S in such a way that each vertex has at most one outgoing edge with a given label. Some subset of the vertices of Γ are called *accept states*. A word w in S^* determines a simplicial path in Γ , by starting at the initial vertex, by reading the letters of w (from left to right) one by one, and by moving along the corresponding edge of Γ if it exists, or halting if not. Associated to Γ there is a subset $L \subset S^*$ consisting of precisely those words that can be read in their entirety without halting, and for which the terminal vertex of the associated path ends at an accept state. One says that L is *parameterized by* (paths in) Γ .

DEFINITION 3.1.1. A subset $L \subset S^*$ is a *regular language* if there is a finite directed graph Γ as above that parameterizes L .

Note that Γ is not part of the data of a regular language, and for any given regular language there will be many graphs that parameterize it. A language is *prefix-closed* if, whenever $w \in L$, every prefix of w is also in L (the empty word is a prefix of every word).

LEMMA 3.1.2. *If L is prefix-closed and regular, there is a Γ parameterizing L for which every vertex is an accept state.*

PROOF. If Γ is any graph that parameterizes L , remove all non-accept vertices and the edges into and out of them. \square

THEOREM 3.1.3 (Generating function). *Let L be a regular language, and for each n , let L_n denote the set of elements of length n , and $L_{\leq n}$ the set of elements of length $\leq n$. Let $s(t) := \sum |L_n|t^n$ and $b(t) := \sum |L_{\leq n}|t^n$ be (formal) generating functions for $|L_n|$ and $|L_{\leq n}|$ respectively. Then $s(t)$ and $b(t)$ are rational; i.e. they agree as power series expansions with some ratio of integral polynomials in t .*

PROOF. Note that $b(t) = s(t)/(1-t)$ so it suffices to prove the theorem for $s(t)$. Let Γ parameterize L , and let M be the adjacency matrix of Γ ; i.e. M_{ij} is equal to the number of edges from vertex i to vertex j . Let v_0 be the vector with a 1 in the initial state, and 0 elsewhere, and let v_a be the vector with a 1 in every accept state, and 0 elsewhere. Then $|L_n| = v_0^T M^n v_a$.

A formal power series $A(t) := \sum a_n t^n$ is rational if and only if its coefficients satisfy a linear recurrence; i.e. if there are constants c_0, \dots, c_d (not all zero) so that $c_0 a_n + c_1 a_{n-1} + \dots + c_d a_{n-d} = 0$ for all $n \geq d$. For, $A(t)(c_0 + c_1 t + \dots + c_d t^d)$ vanishes in degree $\geq d$, and is therefore a polynomial (reversing this argument proves the converse).

If $p(t) = \sum p_i t^{d-i}$ is the characteristic polynomial of M , then $p(M) = 0$, and

$$0 = v_0^T M^{n-d} p(M) v_a = p_0 |L_n| + p_1 |L_{n-1}| + \dots + p_d |L_{n-d}|$$

proving the theorem. \square

Another way of expressing $s(t)$, more useful in some ways, is as follows.

PROPOSITION 3.1.4. *Let L be a regular language. Then there is an integer D so that for each value of $n \bmod D$ either $|L_n|$ is eventually zero, or there are finitely many constants λ_i and polynomials p_i so that $|L_n| = p_1(n)\lambda_1^n + \dots + p_k(n)\lambda_k^n$ for all sufficiently large n .*

For a proof see e.g. [18] Thm. V.3. In particular, either $|L_{\leq n}|$ has polynomial growth, or $C^{-1}(n^k \lambda^n) \leq |L_{\leq n}| \leq C(n^k \lambda^n)$ for some real λ and integer k , and constant C .

3.2. Cannon's theorem. Let S be a set. A total order $<$ on S extends to a unique lexicographic (or dictionary) order on S^* as follows:

- (1) the empty word precedes everything;
- (2) if u and v are both nonempty and start with different letters $s, t \in S$ then $u < v$ if and only if $s < t$; and
- (3) if $u < v$ and w is arbitrary, then $wu < wv$.

If G is a group and S is a generating set for G , there are finitely many geodesic words representing any given element; the lexicographically first geodesic is therefore a canonical representative for each element of G , and determines a language $L \subset S^*$ that bijects with G under evaluation. We denote evaluation by overline, so if $u \in S^*$, we denote the corresponding element of G by \bar{u} . We similarly denote length of an element of S^* by $|\cdot|$. So we always have $|\bar{u}| \leq |u|$ with equality if and only if u is geodesic.

Given $g \in G$ the *cone type* of g , denoted $\text{cone}(g)$, is the set of $h \in G$ for which some geodesic from id to gh passes through g . For any n , the n -level of g is the set of h in the ball $B_n(\text{id})$ such that $|gh| < |g|$. Cannon showed that the n level

(for n sufficiently large) determines the cone type, and therefore that there are only finitely many cone types.

LEMMA 3.2.1 (Cannon [10], Lem. 7.1 p. 139). *The $2\delta + 1$ level of an element determines its cone type.*

PROOF. Let g and h have the same $2\delta + 1$ level, and let u, v be geodesics with $\bar{u} = g$ and $\bar{v} = h$. Only id has an empty $2\delta + 1$ level, so we may assume u, v both have length ≥ 1 . We prove the lemma by induction. Suppose uw, vw and uws are geodesics, where $s \in S$. We must show that vws is a geodesic. Suppose to the contrary that there is some $\bar{w}_1\bar{w}_2 = \bar{vws}$ where $|w_1| = |v| - 1$ and $|w_2| \leq |w| + 1$. Then $h^{-1}\bar{w}_1$ is in the $2\delta + 1$ level of h , which agrees with the $2\delta + 1$ level of g , and therefore $|gh^{-1}\bar{w}_1| < |g|$. But then concatenating a geodesic representative of $gh^{-1}\bar{w}_1$ with w_2 gives a shorter path to \bar{uws} , certifying that uws is not geodesic, contrary to assumption. \square

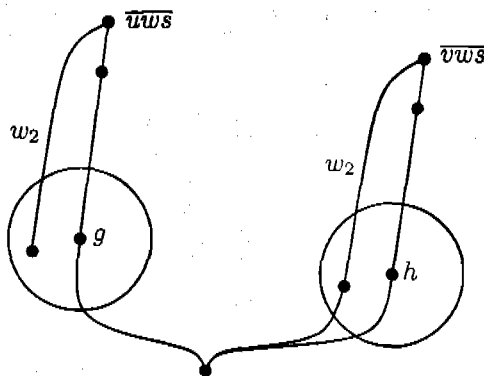


FIGURE 5. A shortcut w_2 from the $2\delta + 1$ level of h to \bar{vws} gives a shortcut from the $2\delta + 1$ level of g to \bar{uws} . This figure is adapted from [15].

The following theorem is implicit in [10], though expressed there in somewhat different language.

THEOREM 3.2.2 (Cannon [10]). *Let G be a hyperbolic group, and S a symmetric generating set. Fix a total order $<$ on S . Then the language of lexicographically first geodesics is prefix-closed and regular.*

PROOF. That this language is prefix-closed is obvious. We show it is regular by describing an explicit parameterizing graph.

As a warm up, we show first that the language of *all* geodesics is regular. A parameterizing graph can be taken as follows. The vertices (all accept states) are precisely the set of cone types, and there is an edge labeled s from a cone type of the form $\text{cone}(g)$ to one of the form $\text{cone}(gs)$ whenever $|gs| = |g| + 1$. By the definition of cone types, this is well-defined. By Lemma 3.2.1, the number of cone types is finite, so this is a finite graph. By construction, this graph exactly parameterizes the language of all geodesics.

Now fix a total order $<$ on S . For each $g \in G$, let u_g be the lexicographically first geodesic from id to g . For each $g \in G$ a *competitor* of g is some h with $|h| = |g|$,

with $u_h \prec u_g$, and for which $d(\overline{u_h|_{\leq i}}, \overline{u_g|_{\leq i}}) \leq 2\delta$ for all i , where $u_g|_{\leq i}$ denotes the prefix of u_g of length i , and similarly for $u_h|_{\leq i}$ (this is described by saying that u_h synchronously fellow-travels u_g).

If there is some g' with $|g'| = |g| + d(g, g')$ and $|g'| = |h| + d(h, g')$ then by δ -thinness and the definition of geodesics, u_h synchronously fellow-travels u_g . It follows that for all $g \in G$ and $s \in S$ we have $u_{gs} = u_g s$ if and only if $u_g s$ is a geodesic, and there is no competitor h of g and $s' \in S$ so that $hs' = gs$.

Given $g \in G$ define $C(g) \subset B_{2\delta}(\text{id})$ to be the set of h for which gh is a competitor of g . Associated to g is the list $L(g)$ of pairs $(h \in C(g), \text{cone}(gh))$ together with the cone type of g itself. Note that the set of possible lists $L(g)$ is finite. We can now define a parameterizing graph by taking the vertices (all accept states) to be the possible lists $L(g)$, and there is an edge labeled s from a list of the form $L(g)$ to a list of the form $L(gs)$ if and only if $|gs| = |g| + 1$, and there is no $h \in C(g)$ and $s' \in S$ with $ghs' = gs$. This is evidently a finite directed graph, which parameterizes the u_g ; we must show it is well-defined.

First of all, $h \in C(gs)$ if and only if one of the two following possibilities occurs:

- (1) there is some $h' \in C(g)$ and $s' \in S \cap \text{cone}(gh')$ with $gh's' = gsh$; or
- (2) there is some $s' \prec s$ in $S \cap \text{cone}(g)$ with $gs' = gsh$.

Both of these possibilities depend only on $C(g)$, $\text{cone}(g)$ or $\text{cone}(gh')$ for some $h' \in C(g)$, and not on g itself. Second of all, if $h \in C(gs)$, then $\text{cone}(gsh)$ depends only on $\text{cone}(gh')$ and s' in the first case, and on $\text{cone}(g)$ and s' in the second case. This shows the graph is well-defined, and completes the proof of the theorem. \square

REMARK 3.2.3. This completes the proof of Lemma 2.5.1, and the subsequent results in § 2.5. The reader will note that the results in this section do not depend on Lemma 2.5.1, so our reasoning has not been circular.

3.3. Combing and combable functions.

DEFINITION 3.3.1. Let G be a group, and S a generating set. A *combing* for G (with respect to S) is a prefix-closed regular language $L \subset S^*$ which bijects with G under evaluation, and satisfies $|\bar{u}| = |u|$ for all $u \in L$ (i.e. L is a language of geodesics).

Theorem 3.2.2 says that every hyperbolic group admits a combing. If L is a combing with respect to S , the L -cone type of g , denoted $\text{cone}_L(g)$, is the set of $h \in G$ for which the L -geodesic evaluating to gh contains a prefix (which is also an L -geodesic) evaluating to g . There is a graph Γ parameterizing L with one vertex for each L -cone type, and an edge from $\text{cone}_L(g)$ to $\text{cone}_L(gs)$ labeled s whenever $s \in \text{cone}_L(g)$.

REMARK 3.3.2. The reader should be warned that many competing definitions of combing exist in the literature.

Suppose L is a combing of G , and Γ is a graph parameterizing L , so that there is a (length-preserving) bijection between directed paths in Γ starting at the initial vertex, and words of L , by reading the edge labels of the path. If $u \in L$, we let $\gamma(u)$ denote the corresponding path in Γ , and $\gamma(u)_i$ the successive vertices in Γ visited by $\gamma(u)$.

DEFINITION 3.3.3. A function $\phi : G \rightarrow \mathbb{Z}$ is *weakly combable* with respect to a combing L if there is a graph Γ parameterizing L and a function $d\phi$ from the vertices of Γ to \mathbb{Z} so that $\phi(\bar{u}) = \sum_i d\phi(\gamma(u)_i)$ for all $u \in L$.

A function ϕ is *combable* if it is weakly combable with respect to some combing L , and if there is a constant C so that $|\phi(gs) - \phi(g)| \leq C$ for all $g \in G$ and $s \in S$; and it is *bicombable* if it is combable, and further satisfies $|\phi(sg) - \phi(g)| \leq C$.

REMARK 3.3.4. It might be more natural to define a function $d\phi$ on the edges of Γ instead of its vertices; however, associated to any directed graph Γ there is another graph — the *line graph* of Γ — whose vertices are the edges of Γ , and whose edges are the composable pairs of edges of Γ , and the line graph of Γ parameterizes L if Γ does.

LEMMA 3.3.5 (Calegari–Fujiwara [8], Lem. 3.8). *The property of being combable or bicombable does not depend on the choice of a generating set or a combing.*

The proof proceeds along the same lines as Theorem 3.2.2. The key point is that words in L are (uniformly) quasigeodesic with respect to S' , and therefore stay within a bounded distance of words in L' with the same evaluation. Therefore an automaton reading the letters of an L' -word can keep track of the states of a collection of automata simultaneously reading nearby L -words, and keeping track of how ϕ changes as one goes along. See [8] for details.

EXAMPLE 3.3.6. Word length in any generating set is bicombable. In fact, if S is a (possibly unsymmetric) set which generates G as a semigroup, word length in S is bicombable. One can generalize word length by giving different generators (and corresponding edges in the Cayley graph) different lengths; providing the lengths are all integral and positive, the resulting (geodesic) word length is bicombable.

EXAMPLE 3.3.7. The sum or difference of two (bi)combable functions is (bi)combable.

EXAMPLE 3.3.8. The following definition is due to Epstein–Fujiwara [16] (also see [5]). Let σ be a path in $C_S(G)$. A *copy* of σ is a translate $g\sigma$ for some $g \in G$. Given a path γ in $C_S(G)$, define $c_\sigma(\gamma)$ to be the maximal number of disjoint copies of σ in γ , and for $g \in G$ define the *small counting function* $c_\sigma : G \rightarrow \mathbb{Z}$ by the formula

$$c_\sigma(g) = |g| - \inf_{\gamma} (|\gamma| - c_\sigma(\gamma))$$

Counting functions are bicombable. In fact, we can add σ to S as a (semigroup) generator, but insist that the (directed) edges labeled σ have length $|\sigma| - 1$ instead of 1; this defines a new distance function $|\cdot|_\sigma$ which is bicombable (by Example 3.3.6), and therefore so is the difference $|\cdot| - |\cdot|_\sigma = c_\sigma$ (by Example 3.3.7).

Many variations on this idea are possible; for instance, the “big” counting functions C_σ which count *all* copies of σ in γ , not just the maximal number of disjoint copies.

3.4. Markov chains. A directed graph Γ is sometimes called a *topological Markov chain*. A topological Markov chain can be promoted to a genuine (stationary) Markov chain by assigning probabilities to each edge in such a way that the probabilities on the edges leaving each vertex sum to 1. Recall that we write the adjacency matrix as M ; we think of this as an endomorphism of the vector space

V spanned by the states of Γ . Let $\mathbf{1}$ denote the vector with all components equal to 1, and let ι denote the vector corresponding to the initial state.

Two states in a topological Markov chain are said to be *communicating* if there is a directed path from each to the other. The property of being communicating is an equivalence relation. We write $C_1 \rightarrow C_2$ for equivalence classes C_1 if there is a directed path from some (any) vertex of C_1 to some (any) vertex of C_2 ; observe that \rightarrow is a partial order. We call each equivalence class a *component*.

The induced (directed) subgraph associated to a component C is itself a topological Markov chain. Its adjacency matrix M_C has the property that for any i and j there is an n (in fact, infinitely many n) so that $(M_C^n)_{ij}$ is positive; one says such a Markov chain is *irreducible*. If there is a fixed n so that $(M_C^n)_{ij}$ is positive for all i, j we say the Markov chain is *aperiodic*; this holds exactly when the gcd of the lengths of all loops in C is 1. A Markov chain (on a finite state space) which is both irreducible and aperiodic is *ergodic*.

LEMMA 3.4.1 (Perron–Frobenius). *Let M be a real matrix with positive entries. Then there is a unique eigenvalue λ of biggest absolute value, and this eigenvalue is real and positive. Moreover, λ is a simple root of the characteristic polynomial, and it has a right (left) eigenvector with all components positive, unique up to scale. Finally, any other non-negative right (left) eigenvector is a multiple of the λ eigenvector.*

PROOF. Since the entries of M are positive, M takes the positive orthant strictly inside itself. The projectivization of the positive orthant is a simplex, and therefore M takes this simplex strictly into its interior. It follows that M has a unique attracting fixed point in the interior of this simplex; this fixed point corresponds to the unique eigenvector v (up to scale) with non-negative entries, and its entries are evidently all positive, and its associated eigenvalue λ is real and positive.

If π is any plane containing this unique positive eigenvector, the projectivization of π is an \mathbb{RP}^1 ; since the eigenvector becomes an attracting fixed point in this \mathbb{RP}^1 , it is not the only fixed point. This shows that λ is a simple eigenvalue; a similar argument shows that $-\lambda$ is not an eigenvalue.

Let μ be any other eigenvalue. If μ is real, then $|\mu| < \lambda$. Suppose μ is complex, acting as composition of a dilation with a rotation on some plane π . If $|\mu| = \lambda$ then the restriction of M to $\pi \oplus \langle v \rangle$ acts projectively like a rotation; but this contradicts the fact that v is a projective attracting fixed point. This proves the theorem. \square

If M is non-negative, there is still a non-negative real eigenvector v with a real positive eigenvalue λ , and every other eigenvalue μ satisfies $|\mu| \leq \lambda$. In this generality, λ might have multiplicity > 1 , and the Jordan block associated to λ might not be diagonal. However if M is *irreducible*, then λ has multiplicity 1, the eigenvector v is strictly positive, and every other eigenvalue with absolute value λ is simple and of the form $e^{2\pi i/k} \lambda$. These facts can be proved similarly to the proof of Lemma 3.4.1.

Now let G be a hyperbolic group, L a combing with respect to some generating set, and Γ a graph parameterizing L . Let Γ_C be the quotient directed graph whose vertices are the components of Γ . Note that Γ_C contains no directed loops. Associated to each vertex of Γ_C is an adjacency matrix M_C which has a unique

maximal real eigenvalue $\lambda(C)$ of multiplicity 1. We let $\lambda = \max_C \lambda(C)$, and we call a component *maximal* if $\lambda(C) = \lambda$.

The next lemma is crucial to what follows, and depends on Coornaert's estimate of the growth function (i.e. Corollary 2.5.9).

LEMMA 3.4.2. *The maximal components do not occur in parallel; that is, there is no directed path from any maximal component to a distinct maximal component.*

PROOF. Since there is a directed path from the initial vertex to every other vertex, the number of paths of length n is of the form $p(n)\lambda^n + O(q(n)\xi^n)$ for polynomials p, q and $\xi < \lambda$, where λ is as above. Moreover, the degree of p is one less than the length of the biggest sequence of maximal components $C_0 \rightarrow C_1 \rightarrow \dots \rightarrow C_{\deg(p)}$. The number of paths of length n is equal to the number of elements of G of length n , so Corollary 2.5.9 implies that the degree of p is zero. \square

It follows that all but exponentially few paths γ of length n in γ are entirely contained in one of the maximal components of Γ , except for a prefix and a suffix of length $O(\log(n))$. Consequently, the properties of a "typical" path in Γ can be inferred from the properties of a "typical" path conditioned to lie in a single component.

For any vector v , the limits

$$\rho(v) := \lim_{n \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \lambda^{-i} M^i v, \quad \ell(v) := \lim_{n \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \lambda^{-i} (M^T)^i v$$

exist, and are the projections onto the left and right λ -eigenspaces respectively. Heuristically, $\ell(v)$ is the distribution of endpoints of long paths that start with distribution v , and $\rho(v)$ is the distribution of starting points of long paths that end with distribution v .

Recall that ι denotes the vector with a 1 in the coordinate corresponding to the initial vertex and 0s elsewhere, and $\mathbf{1}$ denotes the vector with all coordinates equal to 1. Define a measure μ' on the vertices of Γ by $\mu'_i = \ell(\iota)_i \rho(\mathbf{1})_i$, and scale μ' to a probability measure μ . Define a matrix N by $N_{ij} = M_{ij} \rho(\mathbf{1})_j / \lambda \rho(\mathbf{1})_i$ if $\rho(\mathbf{1})_i \neq 0$, and define $N_{ij} = \delta_{ij}$ otherwise.

LEMMA 3.4.3. *The matrix N is a stochastic matrix (i.e. it is non-negative, and the rows sum to 1) and preserves the measure μ .*

PROOF. If $\rho(\mathbf{1})_i = 0$ then $\sum_j N_{ij} = 1$ by fiat. Otherwise

$$\sum_j N_{ij} = \sum_j \frac{M_{ij} \rho(\mathbf{1})_j}{\lambda \rho(\mathbf{1})_i} = \frac{(M \rho(\mathbf{1}))_i}{\lambda \rho(\mathbf{1})_i} = 1$$

To see that N preserves μ' (and therefore μ), we calculate

$$\sum_i \mu'_i N_{ij} = \sum_i \ell(\iota)_i \rho(\mathbf{1})_i \frac{M_{ij} \rho(\mathbf{1})_j}{\lambda \rho(\mathbf{1})_i} = \sum_i \frac{\ell(\iota)_i M_{ij}}{\lambda} \rho(\mathbf{1})_j = \ell(\iota)_j \rho(\mathbf{1})_j = \mu'_j$$

\square

In words, μ_i is the probability that a point on a path will be in state i , conditioned on having originated at the initial vertex in the distant past, and conditioned on having a distant future.

3.5. Shift space. For each n let Y_n denote the set of paths in Γ of length n starting at the initial vertex, and let X_n denote the set of *all* paths in Γ of length n . We can naturally identify X_0 with the vertices of Γ .

Restricting to an initial subpath defines an inverse system $\cdots \rightarrow X_n \rightarrow \cdots \rightarrow X_1 \rightarrow X_0$, and the inverse limit X_∞ is the space of (right) infinite paths. Similarly define $Y_\infty \subset X_\infty$. If we give each X_n and Y_n the discrete topology, then X_∞ and Y_∞ are Cantor sets.

If $x := x_0, x_1, \dots$ and $x' := x'_0, x'_1, \dots$ are two elements of X_∞ , we define $(x|x')$ to be the first index at which x and x' differ, and define a metric on X_∞ by setting $d(x, x') = a^{-(x|x')}$ for some $a > 1$ (the notation $(\cdot|\cdot)$ is deliberately intended to suggest a resemblance to the Gromov product). If we like, we can define $\bar{X} = \bigcup_i X_i \cup X_\infty$ and metrize it (as a compact space, in which each X_n sits as a discrete subset) in the same way. Similarly, give Y_∞ the induced metric, and define $\bar{Y} = \bigcup_i Y_i \cup Y_\infty$ likewise.

The shift operator $T : X_\infty \rightarrow X_\infty$ is defined by $(Tx)_i = x_{i+1}$. We define a probability measure μ on each X_n by $\mu(x_0 \cdots x_n) = \mu_{x_0} N_{x_0 x_1} N_{x_1 x_2} \cdots N_{x_{n-1} x_n}$ where μ and N are the measure and stochastic matrix whose properties are given in Lemma 3.4.3. By the definition of an inverse limit, there is a map $X_\infty \rightarrow X_n$ for each n which takes an infinite path to its initial subpath of length n ; the preimages of subsets of the X_n under such maps are a basis for the topology on X_∞ , called *cylinder sets*. The measures μ as above let us define a Borel probability measure μ on X_∞ by first defining it on cylinder sets (note that the definitions of μ on different X_n are compatible) and extending it to all Borel sets in the standard way; Lemma 3.4.3 implies that μ is T -invariant (i.e. $\mu(A) = \mu(T^{-1}(A))$ for all measurable $A \subset X_\infty$).

There is a bijection between Y_n and L_n , and by evaluation with G_n . This map extends continuously to a map $E : \bar{Y} \rightarrow \bar{G}$, by sending $Y_\infty \rightarrow \partial_\infty G$.

LEMMA 3.5.1. *The map $E : \bar{Y} \rightarrow \bar{G}$ is surjective, Lipschitz in the a -metric, and bounded-to-one.*

PROOF. That the map is Lipschitz follows immediately from the definition, and the observation that $(E(y)|E(y')) \leq (y|y') - \delta$ for $y, y' \in \bar{Y}$. The restrictions $E : Y_n \rightarrow G_n$ are all bijections, so we just need to check that $Y_\infty \rightarrow \partial_\infty G$ is surjective and bounded-to-one.

Since E is continuous, \bar{Y} is compact and \bar{G} is Hausdorff, the image is compact. Since the image is dense (because it contains G_n for all n), it is surjective.

Finally, observe that if y and y' are any two points in Y_∞ , and γ, γ' are the associated infinite geodesics in G , then $\gamma \cap \gamma'$ is a compact initial segment, since after they diverge they never meet again (by the definition of a combing). Fix $x \in \partial_\infty G$, let y_i be a finite subset of $E^{-1}(x)$, and let γ_i be the geodesic rays in G corresponding to the y_i . For all but finitely many points p on any γ_i , each γ_j intersects the ball $B_\delta(p)$ *disjointly* from the others. In particular, the number of points in the preimage of any point in $\partial_\infty G$ is bounded by the cardinality of a ball (in G) of radius δ . \square

Recall that in § 2.5 we defined probability measures ν_s on G for each $s > h(G)$. Note that $e^{h(G)} = \lambda$ where λ is as above. For each n we define a probability measure on Y_n (which, by abuse of notation, we call ν_s) by $\nu_s(y) = \nu_s(E(y)\text{cone}_L(E(y)))$ for $y \in Y_n$, and observe that the limit as $s \rightarrow h(G)$ from above (which we denote $\nu(y)$)

exists and depends only on the cone type $\text{cone}_L(E(y))$. Since the Patterson-Sullivan measure ν is supported in $\partial_\infty G$, the measures ν on each Y_n are compatible, thinking of each Y_n as a collection of cylinder sets in Y_∞ , and define a unique probability measure ν on Y_∞ which pushes forward under E to ν on $\partial_\infty G$.

LEMMA 3.5.2. *The measure μ on X_∞ is the limit $\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} T_*^i \nu$.*

PROOF. We give the sketch of a proof. For any y , let L^y be the (regular) language of suffixes of words in L with y as a prefix, and let L_n^y be the subset of L^y of length n . Then $\nu_s(y) = \zeta_G^{-1}(s) \sum_n e^{-s(|y|+n)} |L_n^y|$.

If there is no path from the final state y_n to a maximal component, the growth rate of L^y is strictly less than that of L , and $\nu_s(y) \rightarrow 0$. Otherwise both growth functions are eventually of the form $C\lambda^n$ plus something exponentially small compared to λ^n . Define measures ν_m on Y_n by $\nu_m(y) = \frac{1}{m} \sum_{i=1}^m \lambda^{-(|y|+i)} |L_i^y|$. Then by considering the form of the growth functions of L and L^y , we see that there is a constant C (not depending on y or n) so that $\lim_{m \rightarrow \infty} \nu_m(y) = C\nu(y)$. Scaling ν_m to be a probability measure, we can set $C = 1$.

The proof now follows from the definition of μ, N ; see [8] Lem. 4.19 for details. \square

3.6. Limit theorems. Let ξ_1, ξ_2, \dots be a (stationary) irreducible Markov chain on a finite state space, with stationary measure μ , and let f be a real-valued function on the state space (since this space is finite, there are no additional assumptions on f ; in general we require f to be integrable, and have finite variance). Define $F_n := \sum_{i=1}^n f(\xi_i)$, and $A = \int f d\mu$.

THEOREM 3.6.1 (Markov's central limit theorem). *With notation as above, there is some $\sigma \geq 0$ so that for any $r \leq s$,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(r \leq \frac{F_n - nA}{\sigma \sqrt{n}} \leq s \right) = \frac{1}{\sqrt{2\pi}} \int_r^s e^{-x^2/2} dx$$

Equivalently, there is convergence in probability $n^{-1/2}(F_n - nA) \rightarrow N(0, \sigma)$ where $N(0, \sigma)$ denote the normal distribution with mean 0 and standard deviation σ (in case $\sigma = 0$, we let $N(0, \sigma)$ denote a Dirac mass centered at 0).

Now, each maximal component C as above is a stationary irreducible Markov chain, with stationary measure the conditional measure $\mu|_C$. The measure μ on X_∞ decomposes measurably into the union of (shift-invariant) subspaces $X_\infty(C)$, the subspace of (right) infinite sequences contained in the component C . Consequently, if ϕ is a combable function on G , then for each maximal component C , there are constants $A_C = \int_C d\phi/\mu(C) d\mu$ and σ_C , so that for μ -a.e. $x \in X_\infty(C)$, the random variable $n^{-1/2}(\sum_{i=0}^{n-1} d\phi(x_i) - nA_C)$ converges in probability to $N(0, \sigma_C)$.

By Lemma 3.5.2, for ν -a.e. $y \in Y_\infty$ there is a unique C so that $T^n y \in X_\infty(C)$ for sufficiently big n ; we say that y is *associated* to the component C . Let $Y_\infty(C)$ be the set of y associated to a fixed C . For ν -a.e. $y \in Y_\infty(C)$ we have convergence in probability $n^{1/2}(\sum_{i=0}^{n-1} d\phi(y_i) - nA_C) \rightarrow N(0, \sigma_C)$ (one way to see this is to observe that this is a shift-invariant tail property of y , and use Lemma 3.5.2).

For combable functions, this is the end of the story. It is certainly possible for the constants A_C, σ_C to vary from component to component. But for bicomable ϕ we have the following key lemma:

LEMMA 3.6.2. *Let ϕ be bicomposable. Then there are constants A, σ so that $A_C = A$ and $\sigma_C = \sigma$ for all maximal components C .*

PROOF. Call $y \in Y_\infty$ *typical* if there are constants A_y and σ_y (necessarily unique) so that $n^{1/2}(\sum_{i=0}^{n-1} d\phi(y_i) - nA_y) \rightarrow N(0, \sigma_y)$. For each C we have seen that ν -a.e. $y \in Y_\infty(C)$ is typical with $A_y = A_C$ and $\sigma_y = \sigma_C$.

The map $E : Y_\infty \rightarrow \partial_\infty G$ is finite-to-one, and takes the measure ν on Y_∞ to the Patterson-Sullivan measure ν on $\partial_\infty G$. Hence $E(Y_\infty(C))$ has positive measure for each C . Let $y \in \partial_\infty G$ be typical, and let $\text{id}, g_1, g_2, \dots$ be the associated geodesic sequence of elements in G converging to $E(y)$. Now let g be arbitrary, let y' be any element of Y_∞ with $E(y') = gE(y)$, and let $\text{id}, g'_1, g'_2, \dots$ be the geodesic sequence of elements in G associated to y' . By δ -thinness, $d(g'_i, gg_i)$ is eventually approximately constant, and therefore bounded. Since ϕ is bicomposable, y' is typical, with $A_{y'} = A_y$ and $\sigma_{y'} = \sigma_y$. But the action of G on ν is ergodic for ν , by Corollary 2.5.10, and therefore for any C, C' there are typical $y \in Y_\infty(C)$, $y' \in Y_\infty(C')$ with $A_y = A_C$, $\sigma_y = \sigma_C$ and $A_{y'} = A_{C'}$, $\sigma_{y'} = \sigma_{C'}$, and with $y' = gy$ for some g . This completes the proof. \square

COROLLARY 3.6.3 (Calegari-Fujiwara [8]). *Let G be hyperbolic, and let ϕ be bicomposable. Then there are constants A, σ so that if g_n denotes a random element of G_n (in the ν measure), there is convergence in probability $n^{-1/2}(\phi(g_n) - nA) \rightarrow N(0, \sigma)$.*

Note that A and σ as above are algebraic, and one can estimate from above the degree of the field extension in which they lie from the complexity of Γ .

The uniform measure and the measure ν on G_n are uniformly quasi-equivalent on a large scale, in the sense that there are constants R and C so that for any $g \in G_n$, there is an inequality

$$C^{-1}|B_R(g) \cap G_n|/|G_n| \leq \nu(B_R(g) \cap G_n) \leq C|B_R(g) \cap G_n|/|G_n|$$

It follows that if g_n denotes a random element of G_n (in the uniform measure), the distribution $n^{-1/2}(\phi(g_n) - nA)$ has a tail that decays like $C_1 e^{-C_2 t^2}$.

Since length with respect to one generating set is bicomposable with respect to another, we obtain the following corollary:

COROLLARY 3.6.4. *Let G be hyperbolic, and let S and S' be two finite generating sets for G . There is an algebraic number $\lambda_{S, S'}$ so that if g_n is a random element of S of word length n , then the distribution $n^{-1/2}(|g_n|_{S'} - n\lambda_{S, S'})$ has a tail that decays like $C_1 e^{-C_2 t^2}$ when n is sufficiently large.*

It is a slightly subtle point that $\lambda_{S', S} \geq \lambda_{S, S'}^{-1}$, and the inequality is strict except for essentially trivial cases.

3.7. Thermodynamic formalism. To push these techniques further, we must study classes of functions more general than combable functions, and invoke more sophisticated limit theorems. There is a well-known framework to carry out such analysis, pioneered by Ruelle, Sinai, Bowen, Ratner, Parry etc.; [36] is a standard reference.

The setup is as follows. For simplicity, let M be a $k \times k$ matrix with 0-1 entries for which there is a constant n so that all the entries of M^n are positive (i.e. M is the adjacency matrix of a topological Markov chain with k states which is irreducible

and aperiodic). Let X_∞ be the space of (right) infinite sequences $x := x_0, x_1, \dots$ satisfying $M(x_n, x_{n+1}) = 1$ for all n , and let T be the shift operator on X_∞ . As before, we can metrize X_∞ by $d(x, x') = a^{-(|x| x')}$ for some fixed $a > 1$, and observe that the action of T on X_∞ is *mixing*. This means that for all nonempty open sets $U, V \subset X_\infty$ there is N so that $T^{-n}(U) \cap V$ is nonempty for all $n \geq N$. Note that if M is irreducible but not aperiodic, there is nevertheless a decomposition of X_∞ into D disjoint components which are cycled by T , and such that T^D is mixing on each component, where D is the gcd of the periods of T -invariant sequences.

Let \mathcal{M}_T be the space of T -invariant probability measures on X_∞ . This is a convex, compact subset of the space of all measures in the weak-* topology. It is not hard to show that the topological entropy h of T is equal to the supremum of the measure theoretic entropies $\sup_{\mu \in \mathcal{M}_T} h(\mu)$, and that $h = \log \lambda$ where λ is the Perron-Frobenius eigenvalue of M ; see e.g. [36].

The shift T uniformly expands X_∞ by a factor of a , and therefore if a function on X_∞ is sufficiently regular, it tends to be smoothed out by T . Define T^*f by $T^*f(x) = f \circ Tx$. We would like the iterates $(T^n)^*f$ to have a uniform modulus of continuity; this is achieved precisely by insisting that f be Hölder continuous, that is, that there is some α so that $|f(x) - f(x')| \leq Cd(x, x')^\alpha = Ca^{-\alpha|x| x'}$. The set of functions f on X_∞ , Hölder continuous of exponent α , is a Banach space with respect to the norm $\|f\|_\infty + \|f\|_\alpha$ where $\|f\|_\alpha$ is the least such C so that $|f(x) - f(x')| \leq Cd(x, x')^\alpha$. We denote this Banach space $C^\alpha(X_\infty)$.

DEFINITION 3.7.1. Let f be Hölder continuous on X_∞ . The *pressure* of f , denoted $P(f)$, is $P(f) = \sup_{\mu \in \mathcal{M}_T} (h(\mu) + \int f d\mu)$.

It turns out that the supremum is realized on some invariant measure μ_f of full support, known as the *equilibrium state* (or *Gibbs state*) of f . That is, $P(f) = h(\mu_f) + \int f d\mu_f$. See e.g. [4] Ch. 1 for a proof of this theorem, and of Theorem 3.7.3 below.

DEFINITION 3.7.2. The *Ruelle transfer operator* L_f associated to f is defined by the formula $L_f g(x) = \sum_{Tx'=x} e^{f(x')} g(x')$. Note that L_f acts as a bounded linear operator on $C^\alpha(X_\infty)$.

THEOREM 3.7.3 (Ruelle-Perron-Frobenius [36]). *The operator L_f has a simple positive eigenvalue $e^{P(f)}$ which is strictly maximal in modulus. The essential spectrum is contained in a ball whose radius is strictly less than $e^{P(f)}$, and the rest of the spectrum outside this ball is discrete and consists of genuine eigenvalues.*

There is a strictly positive eigenfunction ψ_f satisfying $L_f \psi_f = e^{P(f)} \psi_f$, and an "eigen probability measure" ν_f satisfying $L_f^ \nu_f = e^{P(f)} \nu_f$, and if we scale ψ_f so that $\int \psi_f d\nu_f = 1$, then the equilibrium state μ_f is equal to $\nu_f \psi_f$.*

REMARK 3.7.4. ν_f can be thought of as a left eigenvector for L_f , and ψ_f as a right eigenvector. When f is identically zero, L_f is basically just the matrix M , and μ_f is basically just μ as constructed in § 3.4.

Pollicott [34] proved a complexified version of the RPF theorem, and showed that $P(f)$ and ψ_f are analytic on an open subset of the complex Banach space $C^\alpha(X_\infty, \mathbb{C})$ which contains a neighborhood of $C^\alpha(X_\infty, \mathbb{R})$ (i.e. of $C^\alpha(X_\infty)$).

Because of the simplicity and analyticity of the maximal eigenvector/value, one can study the *derivatives* of pressure. For simplicity, let $P(t) := P(tf + g)$. Then

we can compute

$$P'(0) = \int f d\mu_g$$

and a further differentiation gives

$$P''(0) = \int f^2 + 2fw'(0)d\mu_g$$

where $w(t) = \psi_{tf+g}$ (suitably normalized).

Now, let $F_n(x) = \sum_{i=0}^{n-1} f(T^i x)$. Then from the definition of the transfer operator, $L_{tf+g}^n(\cdot) = L_g^n(e^{tF_n} \cdot)$, and therefore one obtains

$$nP''(0) = \int F_n^2 + 2F_n w'(0) d\mu_g$$

If we set g to be identically zero, then μ_g is just the equilibrium measure μ from before. If we change f by a constant $f - \int f d\mu$ to have mean 0, then the ergodic theorem shows $(1/n)F_n \rightarrow 0$ μ -a.e. and therefore

$$P''(0) = \lim_{n \rightarrow \infty} \int \frac{1}{n} F_n^2 d\mu$$

It is usual to denote this limiting quantity by σ^2 .

The analyticity of P lets us control the higher moments of F_n in a uniform manner, and therefore by applying Fourier transform, one obtains a central limit theorem $n^{-1/2}F_n \rightarrow N(0, \sigma)$. Better estimates of the rate of convergence can be obtained by studying $P'''(0)$; see [11].

This theorem can be combined with Lemma 3.6.2 to obtain a central limit theorem for certain functions on hyperbolic groups whose (discrete) derivatives along a combing satisfy a suitable Hölder continuity property. Such functions arise naturally for groups acting cocompactly on CAT(K) spaces with $K < 0$, where one wants to compare the intrinsic geometry of the space with the "coarse" geometry of the group.

Let Z be a complete CAT(K) geodesic metric space with $K < 0$, and let G act cocompactly on Z by isometries. Pick a basepoint $z \in Z$, and define a function F on G by $F(g) = d(z, gz)$. Since G is hyperbolic, if we fix a finite generating set S we can choose a geodesic combing L with respect to S as above. Now, for any $s \in S$ define $D_s F(g) = F(g) - F(sg)$. It is straightforward to see from the CAT(K) property that there are constants C and α (depending on K and G) so that $|D_s F(g) - D_s F(h)| \leq C a^{-\alpha(g|h)}$ for all s and all $g, h \in G$.

An element of $\cup X_n$ corresponds to a path in Γ . Reading the edge labels determines a word in the generators (a suffix of some word in L), and by evaluation, an element of G . Let $E: \cup X_n \rightarrow G$ denote this evaluation map (note that this is not injective). We can define a function DF on $\cup X_n$ by $DF(x) = D_s F(E(x))$ where s^{-1} is the label associated to the transition from x_0 to x_1 (we could suggestively write $s = x_1^{-1}x_0$). Evidently, DF extends to a Hölder continuous function on \bar{X} . Furthermore, for each $y \in Y_n$, we have $\sum_{i=0}^{n-1} DF(T^i y) = F(E(y))$.

For each big component C , it follows that ν -a.e. $y \in Y_\infty(C)$ are A_C, σ_C typical (for the function DF) for some A_C, σ_C depending only on C . Since F is Lipschitz on G in the left and right invariant metrics, the argument of Lemma 3.6.2 implies that A_C, σ_C are equal to some common values A, σ , and therefore we obtain the following corollary:

COROLLARY 3.7.5. *Let Z be a complete $\text{CAT}(K)$ geodesic metric space with $K < 0$, and let G act cocompactly on Z by isometries. Pick a basepoint $z \in Z$, and a finite generating set S for G . Then there are constants A and σ so that if g_n is a random element of G_n (in the ν measure), there is convergence in probability $n^{-1/2}(d(z, g_n z) - An) \rightarrow N(0, \sigma)$.*

Evidently, the only properties of the function F we use are that it is Lipschitz in both the left- and right-invariant metrics, and satisfies a Hölder estimate $|D_s F(g) - D_s F(h)| \leq Ca^{-\alpha(g|h)}$ for all s and all $g, h \in G$. Any such function on a hyperbolic group satisfies a central limit theorem analogous to Corollary 3.7.5. For the sake of completeness, therefore, we state this as a theorem:

THEOREM 3.7.6 (Hölder central limit theorem). *Let G be a hyperbolic group, and S a finite generating set for G . Let F be a real-valued function which is Lipschitz in both the left- and right-invariant word metrics on G , and satisfies $|D_s F(g) - D_s F(h)| \leq Ca^{-\alpha(g|h)}$ for all s in S and all $g, h \in G$. Then there are constants A and σ so that if g_n is a random element of G_n (in the ν measure), there is convergence in probability $n^{-1/2}(F(g_n) - An) \rightarrow N(0, \sigma)$.*

REMARK 3.7.7. The idea of using the thermodynamic formalism to study the relationship between distance and word length in cocompact groups of isometries of hyperbolic space is due to Pollicott–Sharp [35]; Corollary 3.7.5 and Theorem 3.7.6 above are simply the result of combining their work with [12] and [8]. Nevertheless, we believe they are new.

4. Random walks

The main references for this section are Kaimanovich [22] and Kaimanovich–Vershik [23]. The theory of random walks is a vast and deep subject, with connections to many different parts of mathematics. Therefore it is necessary at a few points to appeal to some standard (but deep) results in probability theory, whose proof lies outside the scope of this survey. A basic reference for probability theory is [41]. We give more specialized references in the text where relevant.

This section is brief compared to the earlier sections, and is not meant to be comprehensive.

4.1. Random walk. Let G be a group and let μ be a probability measure on G . We further assume that μ is *nondegenerate*; i.e. that the support of μ generates G as a semigroup. An important example is the case where μ is the uniform measure on a symmetric finite generating set S . There are two ways to describe random walk on G determined by μ : as a sequence of elements visited in the walk, or as a sequence of increments. In the first description, a random walk $y := \text{id}, y_1, y_2, \dots$ is a Markov chain with state space G , with initial state id , and with transition probability $p_{gh} = \mu(g^{-1}h)$. In the second description, a random walk $z := z_1, z_2, \dots$ is a sequence of random elements of G (the increments of the walk), independently distributed according to μ . The two descriptions are related by taking $y_n = z_1 z_2 \dots z_n$. We write this suggestively as $z = Dy$ and $y = \Sigma z$.

We use the notation $(G^{\mathbb{N}}, \mu^{\mathbb{N}})$ for the product probability space, and $(G^{\mathbb{N}}, \mathbf{P})$ for the probability space of infinite sequences with the measure \mathbf{P} on cylinder sets defined by

$$\mathbf{P}(\{y : y \text{ begins } \text{id}, y_1, \dots, y_n\}) = p_{\text{id}y_1} p_{y_1 y_2} \dots p_{y_{n-1} y_n}$$

With this notation, z is a random element of $(G^{\mathbb{N}}, \mu^{\mathbb{N}})$ and y is a random element of $(G^{\mathbb{N}}, \mathbf{P})$.

The shift operator T acts on $G^{\mathbb{N}}$ by $(Tz)_n = z_{n+1}$ or $(Ty)_n = y_{n+1}$. It is measure preserving for $\mu^{\mathbb{N}}$ but not for \mathbf{P} ; in fact, from the definition, the support of \mathbf{P} is contained in the set of sequences starting at id. The action of the shift T on $(G^{\mathbb{N}}, \mu^{\mathbb{N}})$ is ergodic. For, if A is a subset satisfying $A = T^{-1}(A)$, then a sequence z is in A if and only if $T^n(z)$ is in A for sufficiently big n . This is a tail event for the sequence of independent random variables z_i , so by Kolmogorov's 0-1 law (see [41] Thm. 1.1.2) A has measure 0 or 1.

DEFINITION 4.1.1. Let G be a group and S a finite generating set. Let μ be a probability measure on G . The first moment of μ is $\sum |g| \mu(g)$; if this is finite, we say μ has finite first moment.

LEMMA 4.1.2. Let μ be a probability measure on G with finite first moment. Let $\text{id}, y_1, y_2, \dots$ be a random walk determined by μ . Then $L := \lim_{n \rightarrow \infty} |y_n|/n$ exists almost surely, and is independent of y . In fact, if μ^{*n} denotes n -fold convolution (i.e. the distribution of the random variable y_n), then $L = \lim_{n \rightarrow \infty} \sum |g| \mu^{*n}(g)$.

PROOF. We set $z = Dy$. Define $h_n(z) := |y_n|$. Then h_n satisfies

$$h_{n+m}(z) \leq h_n(T^m z) + h_m(z)$$

i.e. h_n form a subadditive cocycle. Kingman's subadditive ergodic theorem (see e.g. [40]) says that for any subadditive L^1 cocycle h_n on a space with a T -invariant measure, the limit $\lim_{n \rightarrow \infty} h_n(z)/n$ exists a.s. and is T -invariant. In our circumstance, finite first moment implies that h_1 (and all the h_n) are in L^1 , so the theorem applies. Since the action of T on $(G^{\mathbb{N}}, \mu^{\mathbb{N}})$ is ergodic, the limit is independent of z . The lemma follows. \square

L as above is called the drift of the random walk associated to μ . Since each $|y_n| \geq 0$ we necessarily have $L \geq 0$.

EXAMPLE 4.1.3. If $G = \mathbb{Z}^n$ and μ is symmetric (i.e. $\mu(g) = \mu(g^{-1})$ for all g) with finite support, then $L = 0$.

We now focus our attention on the case of hyperbolic groups and simple random walk (i.e. when μ is the uniform measure on a finite symmetric generating set).

LEMMA 4.1.4. let G be a nonelementary hyperbolic group, and let μ be a nondegenerate probability measure on G with finite first moment. Then the drift L of random walk with respect to μ is positive.

PROOF. We give the idea of a proof. Let μ^{*n} denote the n -fold convolution of μ as before. The probability measures μ^{*n} have a subsequence converging to a weak limit $\mu^{*\infty}$ in \overline{G} . Clearly the support of $\mu^{*\infty}$ is contained in $\partial_{\infty} G$ (a group for which $\limsup_{n \rightarrow \infty} \mu^{*n}(g) > 0$ for any g and for μ nondegenerate is a finite group).

To prove the lemma it suffices to show that for any C , for sufficiently large enough n there is an inequality $\sum_{g,h} (|hg| - |h|) \mu^{*n}(g) \mu^{*N}(h) \geq C > 0$ for all $N \geq n$, since then $L \geq C/n$. Now, for each h , if g satisfies $|hg| - |h| < C$, then the closest point on the geodesic from h^{-1} to g is within δ of some geodesic from id to h^{-1} . So as $|g|$ goes to infinity, the a -distance from h^{-1} to g goes to 0. Hence for this inequality to fail to hold, almost half of the mass of $\mu^{*n} \times \mu^{*N}$ must be concentrated near the antidiagonal; i.e. the set of $(g, g^{-1}) \in G \times G$.

From this we can deduce that either the desired inequality is satisfied, or else most of the mass of μ^{*n} must be concentrated near a single geodesic through id . Taking $n \rightarrow \infty$, the support of $\mu^{*\infty}$ must consist of exactly two points, and G is seen to be elementary, contrary to hypothesis. \square

REMARK 4.1.5. It is a theorem of Guivarc'h (see [46], Thm. 8.14) that if G is any group with a nondegenerate measure μ (always with finite first moment) for which the drift of random walk is zero, then G is amenable. Some care is required to parse this statement: on an amenable group some nondegenerate measures *may* have positive drift, but on a nonamenable group, *every* nondegenerate measure has positive drift.

A nonelementary hyperbolic group always contains many nonabelian free groups, and is therefore nonamenable; this gives a more highbrow proof of Lemma 4.1.4.

LEMMA 4.1.6 (Kaimanovich [22] 7.2). *Let X be a δ -hyperbolic space. The following two conditions are equivalent for a sequence x_n in X and a number $L > 0$:*

- (1) $d(x_n, x_{n+1}) \leq o(n)$ and $d(x_0, x_n) = nL + o(n)$;
- (2) *there is a geodesic ray γ so that $d(x_n, \gamma(Ln)) = o(n)$.*

A sequence x_n satisfying either condition is said to be *regular*.

PROOF. That (2) implies (1) is obvious, so we show that (1) implies (2). For simplicity, we use the notation $|y| := d(x_0, y)$. The path obtained by concatenating geodesics from x_n to x_{n+1} has finite a -length, and therefore converges to some unique $x_\infty \in \partial_\infty X$.

Let γ_n (resp. γ_∞) be geodesic rays from the origin to x_n (resp. x_∞) and parameterize them by distance from the origin. Fix some positive ϵ , and let $N = N(\epsilon)$ be such that for any two $n, m > N$ the geodesics γ_n and γ_m are within δ on the interval of length $(L - \epsilon)n$, and let $p_n = \gamma_n((L - \epsilon)n)$ so that $d(p_n, \gamma_m) \leq \delta$ for $n > N$. Now, $d(p_{n-1}, p_n) \leq L - \epsilon + 4\delta$ and therefore $d(p_n, p_m) \leq |n - m|(L - \epsilon + 4\delta)$. On the other hand, $d(p_n, p_m) \geq ||p_m| - |p_n|| = |n - m|(L - \epsilon)$. Consequently the sequence p_i is a quasigeodesic, and therefore there is a constant $H = H(\delta, L)$ so that $d(p_n, p_N x_\infty) \leq H$ for any $n \geq N$. Since $p_N x_\infty$ and γ_∞ are asymptotic, $d(p_n, \gamma_\infty) \leq H + \delta$ for sufficiently large n , and therefore $d(x_n, \gamma_\infty) \leq H + \delta + (|x_n| - n(L - \epsilon))$ for sufficiently large n . Taking $\epsilon \rightarrow 0$ proves the lemma with $\gamma = \gamma_\infty$. \square

Together with Lemma 4.1.4 this gives the following Corollary:

COROLLARY 4.1.7 (Kaimanovich [22], 7.3). *Let G be a nonelementary hyperbolic group, and let μ be a nondegenerate probability measure on G with finite first moment. Then there is $L > 0$ so that for a.e. random walk y there is a unique geodesic ray γ_y with $d(y_n, \gamma_y(Ln)) = o(n)$.*

PROOF. It suffices to show that if μ has finite first moment, then $d(y_n, y_{n+1}) = o(n)$ almost surely. Let $z = Dy$, and for any $\epsilon > 0$ let E_n be the event that $|z_n| \geq \epsilon n$. Then the probability of E_n is $\sum_{|g| \geq \epsilon n} \mu(g)$, and therefore

$$\sum_n \mathbb{P}(E_n) = \sum_n \sum_{|g| \geq \epsilon n} \mu(g) \leq \frac{1}{\epsilon} \sum (|g| + 1) \mu(g) < \infty$$

Therefore by the easy direction of the Borel-Cantelli lemma (see e.g. [41] 1.1.4) the probability that E_n occurs infinitely often is zero. Since this is true for every ϵ , we have $d(y_n, y_{n+1}) = |z_n| = o(n)$ almost surely. \square

4.2. Poisson boundary. Define an equivalence relation \sim on $G^{\mathbb{N}}$ by $y \sim y'$ if and only if there are integers k, k' so that $T^k y = T^{k'} y'$.

DEFINITION 4.2.1. The *measurable envelope* of \sim is the smallest measurable equivalence relation generated by \sim . The quotient measure space (Γ, ν) of $(G^{\mathbb{N}}, \mathbf{P})$ by the measurable envelope is called the *Poisson boundary* of G with respect to μ .

In other words, ν -measurable functions on Γ correspond precisely to T -invariant \mathbf{P} -measurable functions on $G^{\mathbb{N}}$. We let $\text{bnd} : G^{\mathbb{N}} \rightarrow \Gamma$ be the quotient map, so that $\text{bnd } \mathbf{P} = \nu$.

Now, G acts on $G^{\mathbb{N}}$ on the left coordinatewise. This action commutes with T , and descends to an action on Γ . Since \sim is T -invariant, $\text{bnd } \mathbf{P} = \text{bnd } T\mathbf{P}$, so $\nu = \sum_g \mu(g)g\nu$; i.e. the measure ν is μ -stationary.

DEFINITION 4.2.2. A μ -boundary is a G -space with a μ -stationary measure λ which is obtained as a T -equivariant (measurable) quotient of $(G^{\mathbb{N}}, \mathbf{P})$.

Any μ -boundary factors through (Γ, ν) . A μ -boundary is μ -maximal if the map from (Γ, ν) is a measurable isomorphism. Kaimanovich [22] gave two very useful criteria for a μ -boundary to be maximal.

THEOREM 4.2.3 (Kaimanovich ray criterion [22], Thm. 5.5). *Let B be a μ -boundary, and for $y \in G^{\mathbb{N}}$ let $\Pi(y) \in B$ be the image of y under the (G -equivariant) quotient map $\Pi : G^{\mathbb{N}} \rightarrow B$. If there is a family of measurable maps $\pi_n : B \rightarrow G$ such that \mathbf{P} -a.e. $d(y_n, \pi_n(\Pi(y))) = o(n)$ then B is maximal.*

Together with Corollary 4.1.7, this gives the following important result:

COROLLARY 4.2.4 (Kaimanovich [22], Thm. 7.6). *Let G be a nonelementary hyperbolic group, and let μ be a nondegenerate probability measure of finite first moment. Let $\Pi : G^{\mathbb{N}} \rightarrow \partial_{\infty} G$ take a random walk to its endpoint (which exists \mathbf{P} -a.e.), and let $\lambda = \Pi\mathbf{P}$. Then $(\partial_{\infty} G, \lambda)$ is the Poisson boundary of G, μ .*

PROOF. Simply define π_n to be the maps that take a point $y \in \partial_{\infty} G$ to $\gamma_y(nL)$ where γ_y is a parameterized geodesic ray from id to y , and L is the drift. \square

4.3. Harmonic functions.

DEFINITION 4.3.1. If f is a function on G , the operator P_{μ} (convolution with μ) is defined by $P_{\mu}f(g) := \sum_h f(gh)\mu(h)$. A function f on G is μ -harmonic (or just harmonic if μ is understood) if it is fixed by P_{μ} ; i.e. if it satisfies $f(g) = \sum_h f(gh)\mu(h)$ for all g in G .

In general we need to impose some condition on f for $\sum_h f(gh)\mu(h)$ to be defined. If the support of μ is finite, then f can be arbitrary, but if the support of μ is infinite, we usually (but not always!) require f to be in L^{∞} . We let $H^{\infty}(G, \mu)$ denote the Banach space of bounded μ -harmonic functions on G .

In probabilistic terms, if f is harmonic and $y \in G^{\mathbb{N}}$ is a random walk, the random variables $f_n := f(y_n)$ are a *martingale*; i.e. the expected value of f_n given y_{n-1} is f_{n-1} (see e.g. [41] § 5.2 for an introduction to martingales). There is an intimate relation between harmonic functions and Poisson boundaries, expressed in the following proposition.

LEMMA 4.3.2. *The Banach spaces $H^{\infty}(G, \mu)$ and $L^{\infty}(\Gamma, \nu)$ are isometric.*

PROOF. Given $f \in H^\infty(G, \mu)$ and $y \in G^{\mathbb{N}}$, the random variables $f(y_n)$ are a bounded martingale, and therefore by the martingale convergence theorem ([41] Thm. 5.2.22), converge a.s. to a well-defined limit. Evidently this limit is measurable and T -invariant, and therefore descends to a function on Γ which we denote \hat{f} . Explicitly, $\hat{f}(\text{bnd } y) := \lim_{n \rightarrow \infty} f(y_n)$.

Conversely, given $\hat{f} \in L^\infty(\Gamma, \nu)$ we define $f(g) = \int_{\Gamma} \hat{f} d(g_* \nu)$ (this expression is known as the *Poisson formula*). Since ν is stationary, f is harmonic.

The mean value property of harmonic functions implies that these maps are isometries, since a harmonic function achieves its maximum on the boundary. \square

Note that the Poisson formula is available for *any* μ -boundary. That is, if B is a G -space with a μ -stationary probability measure λ , and \hat{f} is any element of $L^\infty(B, \lambda)$, then $f(g) := \int_B \hat{f} d(g\lambda)$ is a bounded harmonic function on G . If λ is not invariant, f is typically nonconstant.

The remainder of this section is devoted to some miscellaneous applications of random walks to hyperbolic and other groups.

4.4. Green metric. There is a close resemblance between the measure ν and the Patterson–Sullivan measures constructed in § 2.5. This resemblance can be sharpened if one looks at a natural metric on G adapted to the random walk, namely the so-called *Green metric*.

DEFINITION 4.4.1. Let G be a group and μ a probability measure on G with finite first moment. The *Green metric* on G is the metric for which the distance between g and h is $-\log$ of the probability that random walk starting at g ever hits h .

If μ is symmetric, so is the Green metric, since random walks are time-reversible. Note that the Green metric is degenerate if random walk is recurrent. For simple random walk, this occurs only if G is finite, or is virtually \mathbb{Z} or \mathbb{Z}^2 , by a classical result of Varopoulos (see [45]). For nondegenerate measures with finite first moment on non-elementary hyperbolic groups, Blachère and Brofferio [1] show that the Green metric and the word metric are quasi-isometric (one needs to be somewhat careful: the Green metric is *not* in general a geodesic metric).

THEOREM 4.4.2 (Blachère–Haïssinsky–Mathieu [2], Thm. 1.3). Let G be a non-elementary hyperbolic group, and for $y \in \partial_\infty G$, let $B(y, R)$ denote the ball of radius R in the a -metric (see Definition 2.4.6). Let μ be a symmetric probability measure with finite first moment, and let ν be the associated harmonic measure on $\partial_\infty G$. Then for ν -almost every $y \in \partial_\infty G$, there is convergence

$$\lim_{R \rightarrow 0} \log \nu(B(y, R)) / \log R = \ell_G / aL$$

where L is the drift in the word metric, and ℓ_G is the drift in the Green metric.

Note that Kingman's subadditive ergodic theorem implies that the drift ℓ_G with respect to the Green metric is well-defined, essentially by the same argument as the proof of Lemma 4.1.2.

4.5. Harnack inequality. The classical *Harnack inequality* relates the values of a positive harmonic function at two points. In its infinitesimal version, it asserts an upper bound on the logarithmic derivative of a positive harmonic function.

Let f be a non-negative bounded harmonic function on \mathbb{H}^n , for simplicity. The Poisson formula says that $f(p) = \int_{S_\infty^{n-1}} f d\nu_p$ where ν_p is the visual measure as seen from p . If ν is visual measure as seen from the origin, and g is any isometry taking the origin to p , then $\nu_p = g_*\nu$. To understand how f varies as a function of p therefore, it suffices to understand how ν_p varies as a function of p . If B is an infinitesimal ball centered at some point y in S_∞^{n-1} , then the visual size of B grows like $e^{t(n-1)}$ as one moves distance t in the direction of y . Hence:

PROPOSITION 4.5.1 (Harnack inequality). *Let f be a non-negative bounded harmonic function on \mathbb{H}^n . Then the logarithmic derivative of f satisfies the inequality $|d \log f| \leq (n-1)$.*

If f is a non-negative harmonic function on a group G , the analog of this inequality is $f(gs)/f(g) \leq e^D$ for any $g \in G$ and $s \in S$ where D is the dimension of ν , which can be determined from Theorem 4.4.2.

If S is a closed surface of genus ≥ 2 and $\rho: \pi_1(S) \rightarrow G$ is injective, then $\pi_1(S)$ acts on G by left translation, and there is an associated foliated bundle with fiber the ideal circle $\partial_\infty \pi_1(S)$ with its natural $\pi_1(S)$ action. We can build a harmonic connection for this circle bundle; i.e. a choice of measure m_g on the circle $S^1(g)$ over each $g \in G$ so that for any subset $A \subset S^1$ we have $m_g(A) = \sum \mu(s) m_{gs}(A)$. Since the circle is 1-dimensional, these measures integrate to *metrics* on the circles $S^1(g)$ for which the curvature is harmonic. The Harnack inequality then gives *a priori* bounds on this curvature, and one can deduce local compactness results for families of injective surface maps of variable genus. For stable minimal surfaces in hyperbolic 3-manifolds, such *a priori* bounds were obtained by Schoen [38] and are an important tool in low-dimensional topology. The idea of using Harnack-type inequalities to obtain curvature bounds is due to Thurston [44] (also see [6], Example 4.6).

4.6. Monotonicity. A *norm* on a group is a non-negative function $\tau: G \rightarrow \mathbb{R}$ so that $\tau(gh) \leq \tau(g) + \tau(h)$ for all $g, h \in G$. A functor from groups to norms is *monotone* if $\tau_H(\phi(g)) \leq \tau_G(g)$ for any $g \in G$ and $\phi: G \rightarrow H$.

If τ is a norm on G , and μ is a probability measure with finite first moment, it makes sense to study the growth rate of τ under μ -random walk on G . If G is finitely generated, one can study the growth rate of τ under all simple random walks; if they all have the same growth rate, this rate is an invariant of G . Since μ random walk on G pushes forward to $\phi_*\mu$ random walk on $\phi(G) = H$, the growth rate of a monotone family of norms cannot increase under a homomorphism; thus if the growth rate of τ_G on G is strictly smaller than the growth rate of τ_H on H , there are strong constraints on the homomorphisms from G to H .

As an example, consider the *commutator length* cl . For any group G and any g in the commutator subgroup $[G, G]$, the commutator length $cl(g)$ is just the least number of commutators in G whose product is g (for technical reasons, one usually studies a closely related quantity, namely the *stable commutator length*; see e.g. [7] for an introduction).

One of the main theorems of [9] is as follows:

THEOREM 4.6.1 (Calegari-Maher [9]). *Let G be hyperbolic, and let μ be a nondegenerate symmetric probability measure with finite first moment whose support generates a nonelementary subgroup. There is a constant C so that if g_n is obtained*

by random walk of length n , conditioned to lie in $[G, G]$, then

$$C^{-1}n/\log(n) \leq \text{cl}(g_n) \leq Cn/\log(n)$$

with probability $1 - O(C^{-n^\epsilon})$.

Said another way, commutator length grows like $n/\log(n)$ under random walk in a hyperbolic group. Similar estimates on commutator length can be obtained for groups acting in a suitable way on (not necessarily proper) hyperbolic spaces; the most important examples are mapping class groups and relatively hyperbolic groups.

As a corollary, if H is any finitely generated group, and commutator length in H grows like $o(n/\log(n))$ for simple random walk (with respect to some generating set), then there are no interesting homomorphisms from H to any hyperbolic group G , and no interesting actions of H on certain hyperbolic complexes.

Acknowledgments

I would like to thank Vadim Kaimanovich, Anders Karlsson, Joseph Maher, Curt McMullen, Richard Sharp, and Alden Walker. I would also like to thank the anonymous referee for some useful comments. Danny Calegari was supported by NSF grant DMS 1005246.

References

- [1] Sébastien Blachère and Sara Brofferio, *Internal diffusion limited aggregation on discrete groups having exponential growth*, Probab. Theory Related Fields **137** (2007), no. 3-4, 323–343, DOI 10.1007/s00440-006-0009-2. MR2278460 (2008b:60210)
- [2] Sébastien Blachère, Peter Haüssinsky, and Pierre Mathieu, *Harmonic measures versus quasi-conformal measures for hyperbolic groups*, Ann. Sci. Éc. Norm. Supér. (4) **44** (2011), no. 4, 683–721 (English, with English and French summaries). MR2919980
- [3] Brian H. Bowditch, *A topological characterisation of hyperbolic groups*, J. Amer. Math. Soc. **11** (1998), no. 3, 643–667, DOI 10.1090/S0894-0347-98-00264-1. MR1602069 (99c:20048)
- [4] Rufus Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Second revised edition, Lecture Notes in Mathematics, vol. 470, Springer-Verlag, Berlin, 2008. With a preface by David Ruelle; edited by Jean-René Chazottes. MR2423393 (2009d:37038)
- [5] Robert Brooks, *Some remarks on bounded cohomology*, Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), Ann. of Math. Stud., vol. 97, Princeton Univ. Press, Princeton, N.J., 1981, pp. 53–63. MR624804 (83a:57038)
- [6] Danny Calegari, *Foliations and the geometry of 3-manifolds*, Oxford Mathematical Monographs, Oxford University Press, Oxford, 2007. MR2327361 (2008k:57048)
- [7] Danny Calegari, *scl*, MSJ Memoirs, vol. 20, Mathematical Society of Japan, Tokyo, 2009. MR2527432 (2011b:57003)
- [8] Danny Calegari and Koji Fujiwara, *Combable functions, quasimorphisms, and the central limit theorem*, Ergodic Theory Dynam. Systems **30** (2010), no. 5, 1343–1369, DOI 10.1017/S0143385709000662. MR2718897 (2011k:20088)
- [9] D. Calegari and J. Maher, *Statistics and compression of scl*, preprint, arXiv:1008.4952; to appear in Ergodic Theory and Dynamical Systems.
- [10] James W. Cannon, *The combinatorial structure of cocompact discrete hyperbolic groups*, Geom. Dedicata **16** (1984), no. 2, 123–148, DOI 10.1007/BF00146825. MR758901 (86j:20032)
- [11] Zaqueu Coelho and William Parry, *Central limit asymptotics for shifts of finite type*, Israel J. Math. **69** (1990), no. 2, 235–249, DOI 10.1007/BF02937307. MR1045376 (91f:60065)
- [12] Michel Coornaert, *Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov*, Pacific J. Math. **159** (1993), no. 2, 241–270 (French, with French summary). MR1214072 (94m:57075)

- [13] M. Coornaert, T. Delzant, and A. Papadopoulos, *Géométrie et théorie des groupes*, Lecture Notes in Mathematics, vol. 1441, Springer-Verlag, Berlin, 1990 (French). Les groupes hyperboliques de Gromov. [Gromov hyperbolic groups]; With an English summary. MR1075994 (92f:57003)
- [14] Michel Coornaert and Athanase Papadopoulos, *Symbolic dynamics and hyperbolic groups*, Lecture Notes in Mathematics, vol. 1539, Springer-Verlag, Berlin, 1993. MR1222644 (94d:58054)
- [15] David B. A. Epstein, James W. Cannon, Derek F. Holt, Silvio V. F. Levy, Michael S. Paterson, and William P. Thurston, *Word processing in groups*, Jones and Bartlett Publishers, Boston, MA, 1992. MR1161694 (93i:20036)
- [16] David B. A. Epstein and Koji Fujiwara, *The second bounded cohomology of word-hyperbolic groups*, Topology 36 (1997), no. 6, 1275–1289, DOI 10.1016/S0040-9383(96)00046-8. MR1452851 (98k:20088)
- [17] Herbert Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969. MR0257325 (41 #1976)
- [18] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009. MR2483235 (2010h:05005)
- [19] Alex Furman, *Random walks on groups and random transformations*, Handbook of dynamical systems, Vol. 1A, North-Holland, Amsterdam, 2002, pp. 931–1014, DOI 10.1016/S1874-575X(02)80014-5. MR1928529 (2003j:60065)
- [20] M. Gromov, *Hyperbolic groups*, Essays in group theory, Math. Sci. Res. Inst. Publ., vol. 8, Springer, New York, 1987, pp. 75–263, DOI 10.1007/978-1-4613-9586-7.3. MR919829 (89e:20070)
- [21] P. Hall and C. C. Heyde, *Martingale limit theory and its application*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. MR624435 (83a:60001)
- [22] Vadim A. Kaimanovich, *The Poisson formula for groups with hyperbolic properties*, Ann. of Math. (2) 152 (2000), no. 3, 659–692, DOI 10.2307/2661351. MR1815698 (2002d:60064)
- [23] V. A. Kaimanovich and A. M. Vershik, *Random walks on discrete groups: boundary and entropy*, Ann. Probab. 11 (1983), no. 3, 457–490. MR704539 (85d:60024)
- [24] Shizuo Kakutani, *Random ergodic theorems and Markoff processes with a stable distribution*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, University of California Press, Berkeley and Los Angeles, 1951, pp. 247–261. MR0044773 (13,476a)
- [25] Michael Kapovich and Bruce Kleiner, *Hyperbolic groups with low-dimensional boundary*, Ann. Sci. École Norm. Sup. (4) 33 (2000), no. 5, 647–669, DOI 10.1016/S0012-9593(00)01049-1 (English, with English and French summaries). MR1834498 (2002j:20077)
- [26] Harry Kesten, *Full Banach mean values on countable groups*, Math. Scand. 7 (1959), 146–156. MR0112053 (22 #2911)
- [27] Colin Maclachlan and Alan W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics, vol. 219, Springer-Verlag, New York, 2003. MR1937957 (2004i:57021)
- [28] Bernard Maskit, *Kleinian groups*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 287, Springer-Verlag, Berlin, 1988. MR959135 (90a:30132)
- [29] Peter J. Nicholls, *The ergodic theory of discrete groups*, London Mathematical Society Lecture Note Series, vol. 143, Cambridge University Press, Cambridge, 1989. MR1041575 (91i:58104)
- [30] S. J. Patterson, *The limit set of a Fuchsian group*, Acta Math. 136 (1976), no. 3-4, 241–273. MR0450547 (56 #8841)
- [31] G. Perelman, *The entropy formula for the Ricci flow and its geometric applications*, preprint arXiv:math/0211159
- [32] G. Perelman, *Ricci flow with surgery on three-manifolds*, preprint arXiv:math/0303109
- [33] Jean-Claude Picaud, *Cohomologie bornée des surfaces et courants géodésiques*, Bull. Soc. Math. France 125 (1997), no. 1, 115–142 (French, with English and French summaries). MR1459300 (98h:58146)
- [34] Mark Pollicott, *A complex Ruelle-Perron-Frobenius theorem and two counterexamples*, Ergodic Theory Dynam. Systems 4 (1984), no. 1, 135–146, DOI 10.1017/S0143385700002327. MR758899 (88i:58097)

- [35] Mark Pollicott and Richard Sharp, *Comparison theorems and orbit counting in hyperbolic geometry*, Trans. Amer. Math. Soc. **350** (1998), no. 2, 473–499, DOI 10.1090/S0002-9947-98-01756-5. MR1376553 (98d:58146)
- [36] David Ruelle, *Thermodynamic formalism*, Encyclopedia of Mathematics and its Applications, vol. 5, Addison-Wesley Publishing Co., Reading, Mass., 1978. The mathematical structures of classical equilibrium statistical mechanics; with a foreword by Giovanni Gallavotti and Gian-Carlo Rota. MR511655 (80g:82017)
- [37] Richard Sharp, *Local limit theorems for free groups*, Math. Ann. **321** (2001), no. 4, 889–904, DOI 10.1007/s002080100258. MR1872533 (2002k:20039)
- [38] Richard Schoen, *Estimates for stable minimal surfaces in three-dimensional manifolds*, Seminar on minimal submanifolds, Ann. of Math. Stud., vol. 103, Princeton Univ. Press, Princeton, NJ, 1983, pp. 111–126. MR795231 (86j:53094)
- [39] John R. Stallings, *On torsion-free groups with infinitely many ends*, Ann. of Math. (2) **88** (1968), 312–334. MR0228573 (37 #4153)
- [40] J. Michael Steele, *Kingman's subadditive ergodic theorem*, Ann. Inst. H. Poincaré Probab. Statist. **25** (1989), no. 1, 93–98 (English, with French summary). MR995293 (90e:28025)
- [41] Daniel W. Stroock, *Probability theory, an analytic view*, Cambridge University Press, Cambridge, 1993. MR1267569 (95f:60003)
- [42] Dennis Sullivan, *On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions*, Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), Ann. of Math. Stud., vol. 97, Princeton Univ. Press, Princeton, N.J., 1981, pp. 465–496. MR624833 (83f:58052)
- [43] Dennis Sullivan, *The density at infinity of a discrete group of hyperbolic motions*, Inst. Hautes Études Sci. Publ. Math. **50** (1979), 171–202. MR556586 (81b:58031)
- [44] W. Thurston, *3-manifolds, foliations and circles II*, preprint
- [45] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992. MR1218884 (95f:43008)
- [46] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000. MR1743100 (2001k:60006)

DEPARTMENT OF MATHEMATICS, CALTECH, PASADENA, CALIFORNIA 91125
 E-mail address: dannyc@math.uchicago.edu